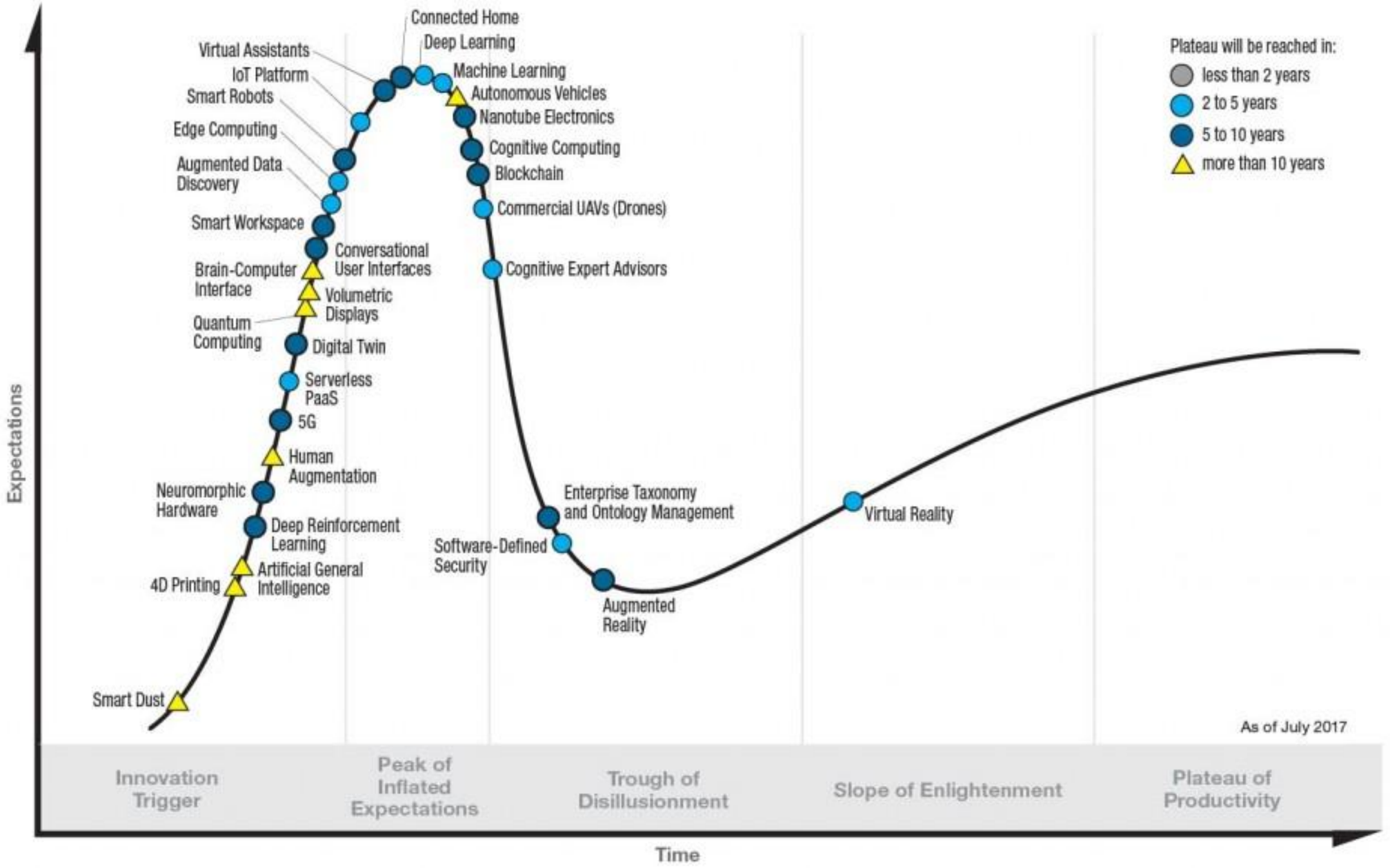
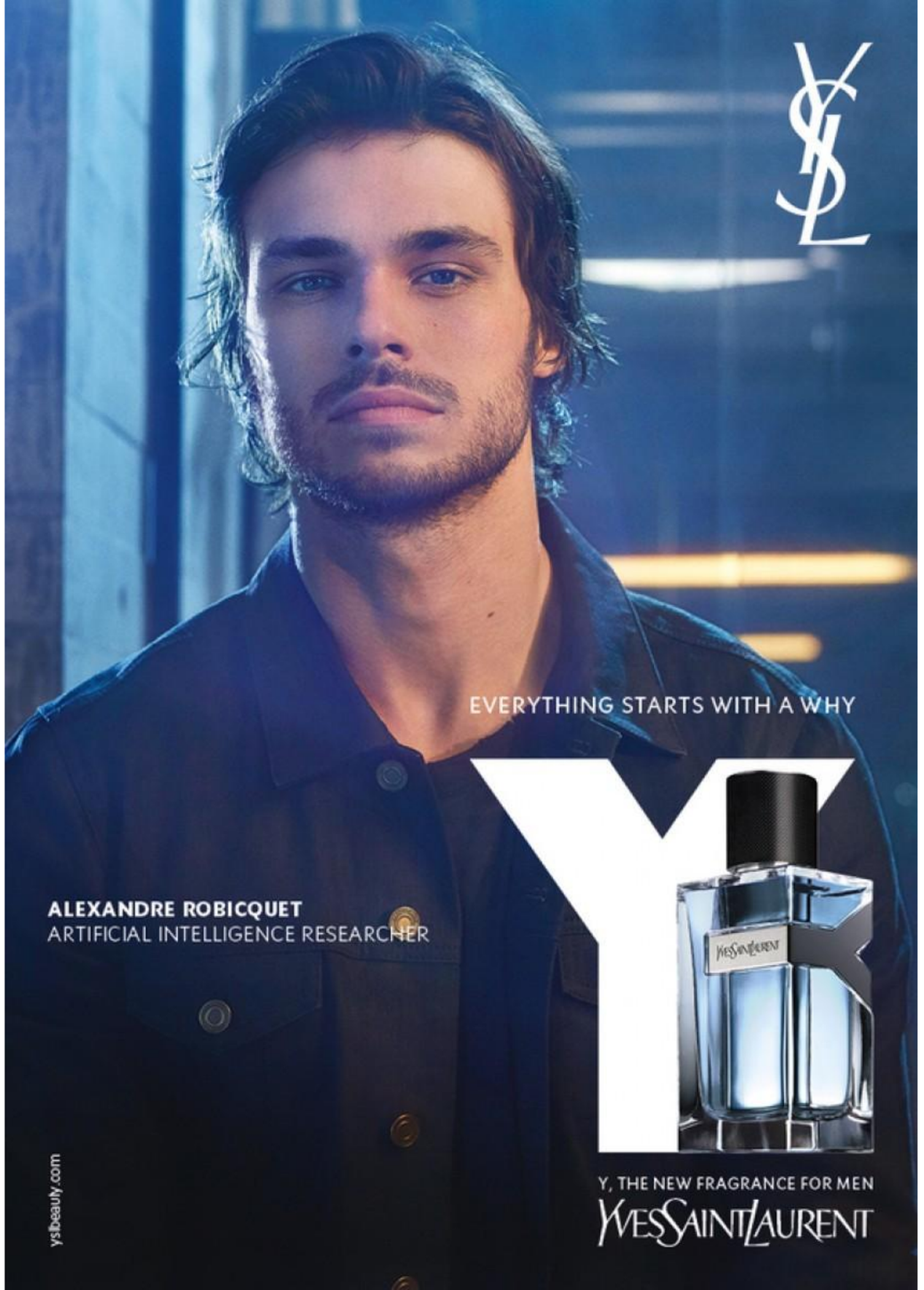
Introduction:
A new approach to ML

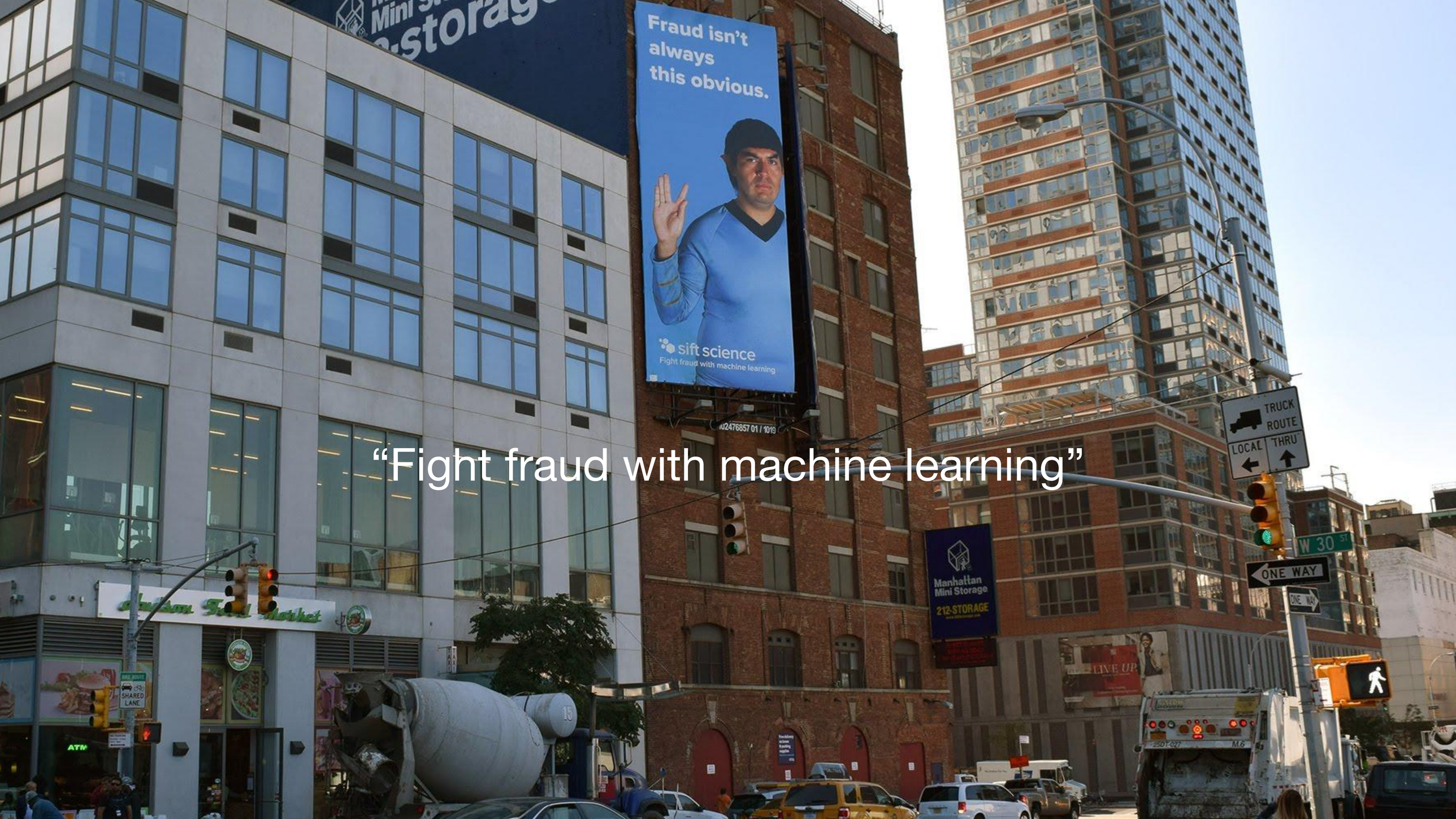
@HumanCoders #ML #workshop

All Rights Reserved © Louis Dorard

Gartner Hype Cycle for Emerging Technologies, 2017




EVERYTHING STARTS WITH A WHY
ALEXANDRE ROBICQUET
ARTIFICIAL INTELLIGENCE RESEARCHER
Y, THE NEW FRAGRANCE FOR MEN
YVES SAINT LAURENT
ysbeauty.com



"Fight fraud with machine learning"

Jian-Yang's actually put together
a pretty good classifier.

Orange F 4G 21:04 48 %

Chocolate Mousse

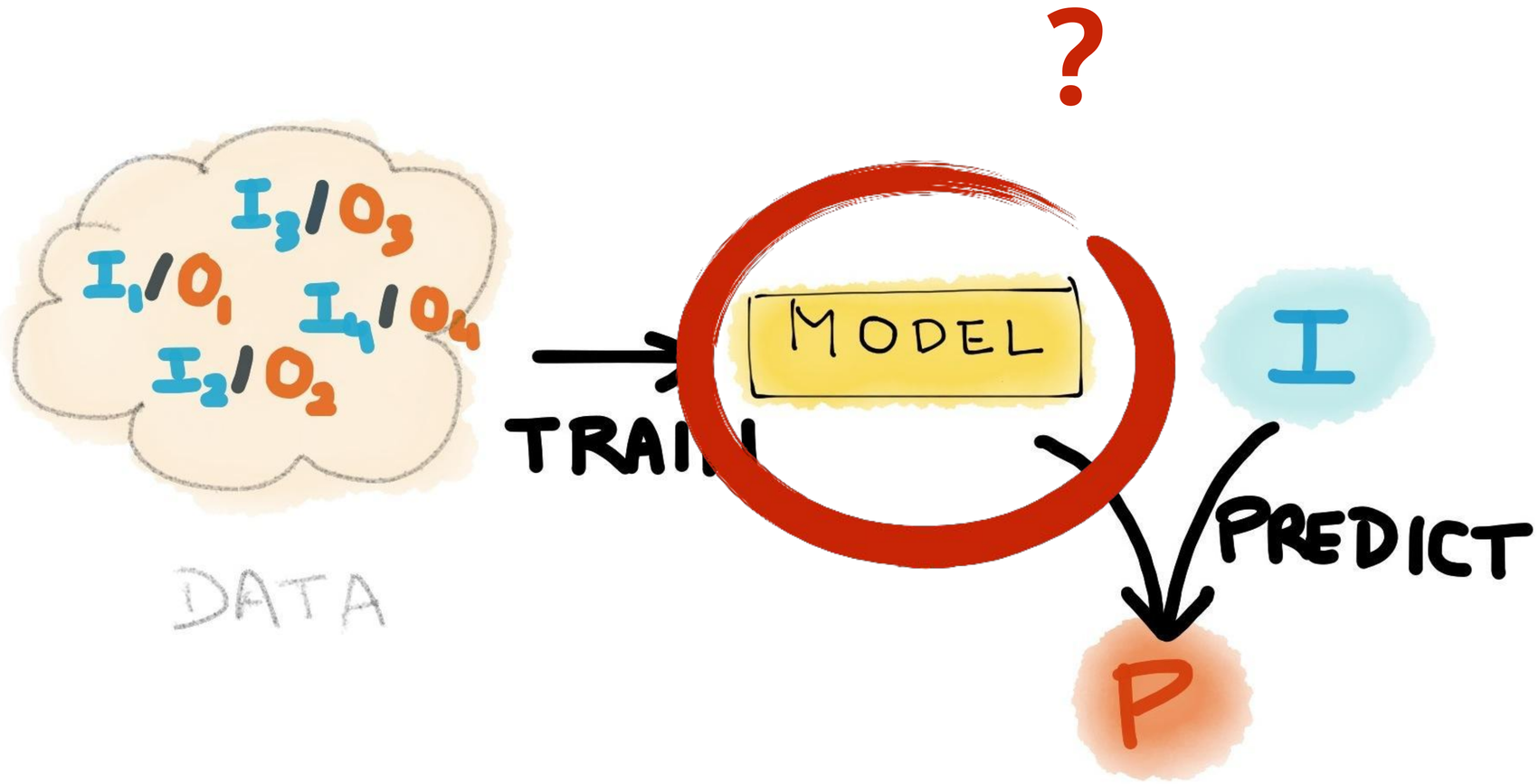
44%

Cals	Fat	Carb	Prot
355	25	27	7

Per 1 cup

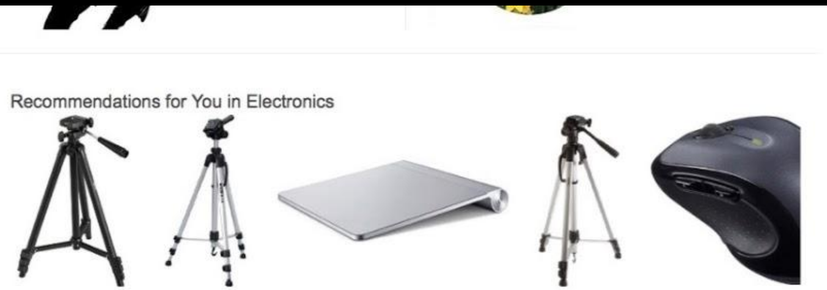
last column = output (by convention)

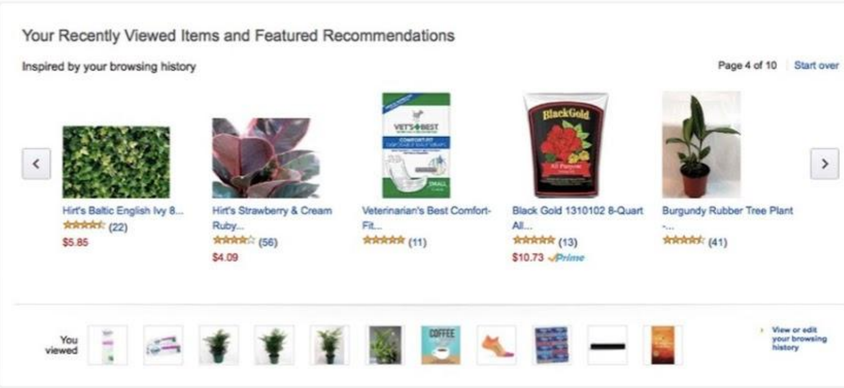
Bedrooms	Bathrooms	Surface (foot²)	Year built	Type	Price ($)
3	1	860	1950	house	565,000
3	1	1012	1951	house	
2	1.5	968	1976	townhouse	447,000
4		1315	1950	house	648,000
3	2	1599	1964	house	
3	2	987	1951	townhouse	790,000
1	1	530	2007	condo	122,000
4	2	1574	1964	house	835,000
4			2001	house	855,000
3	2.5	1472	2005	house	
4	3.5	1714	2005	townhouse	
2	2	1113	1999	condo	
1		769	1999	condo	315,000


?
I3/O3
I1/O1
I4/O4
I2/O2
DATA
MODEL
TRAIN
I
PREDICT
P

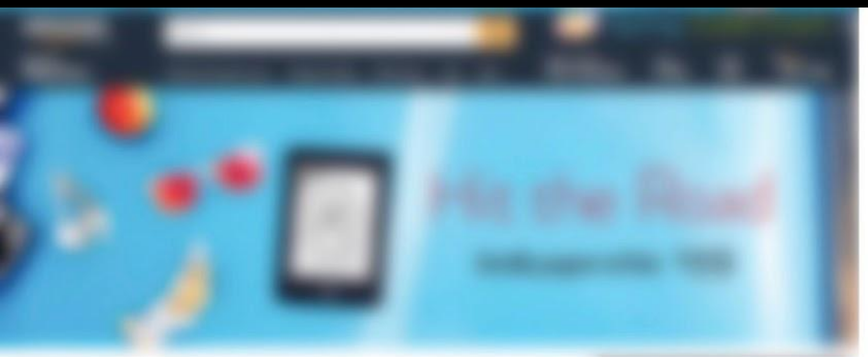
ML and AI are everywhere

Amazon for David Jones (@d_jones, see source)





Amazon for David Jones (@d_jones, see source)



Related to Items You've Viewed See more

Recommendations for You in Health & Personal Care

More Recommendations for You See more

Recommendations for You in Electronics

More Items to Consider See more

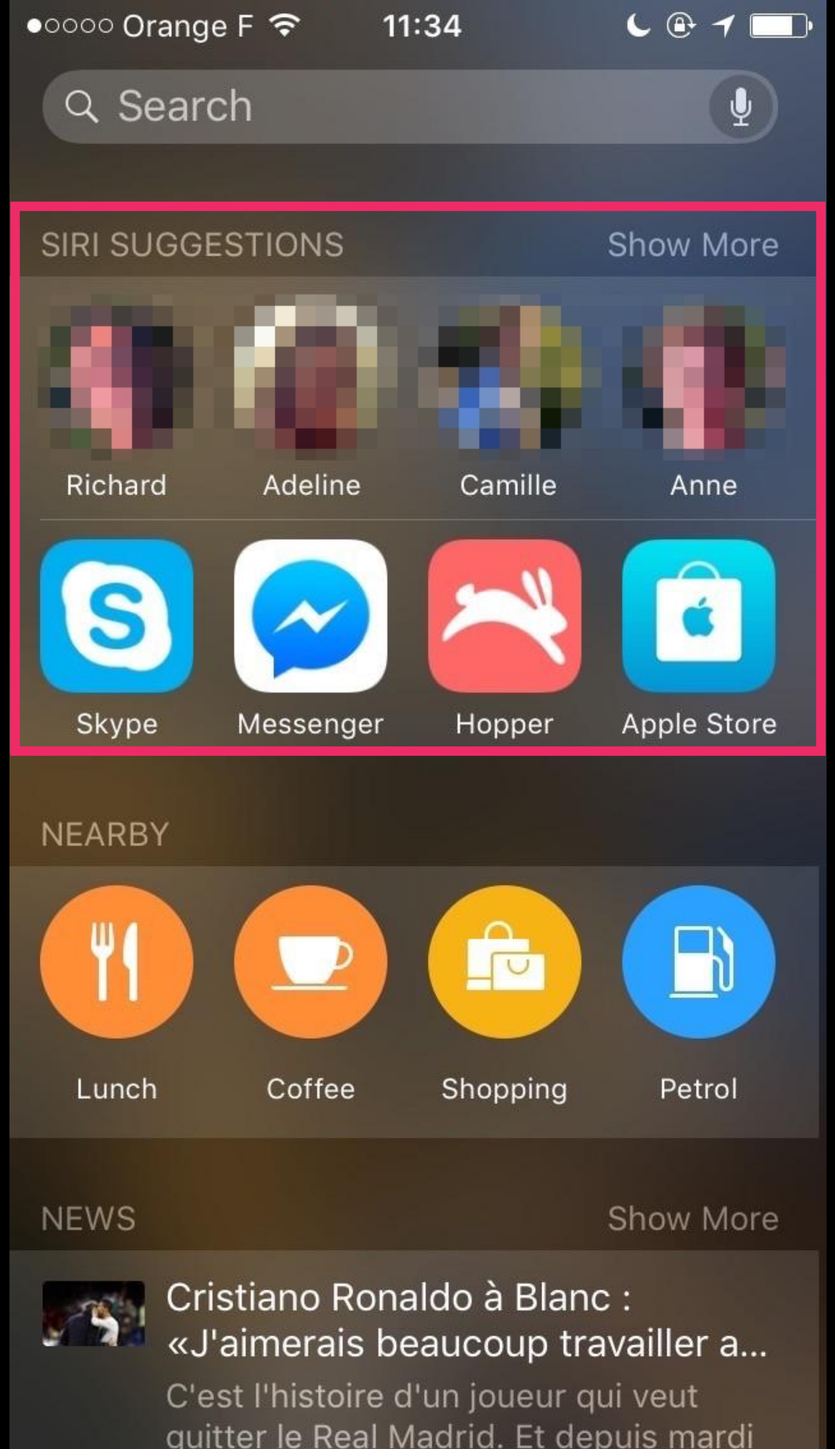
Inspired by Your Browsing History See more

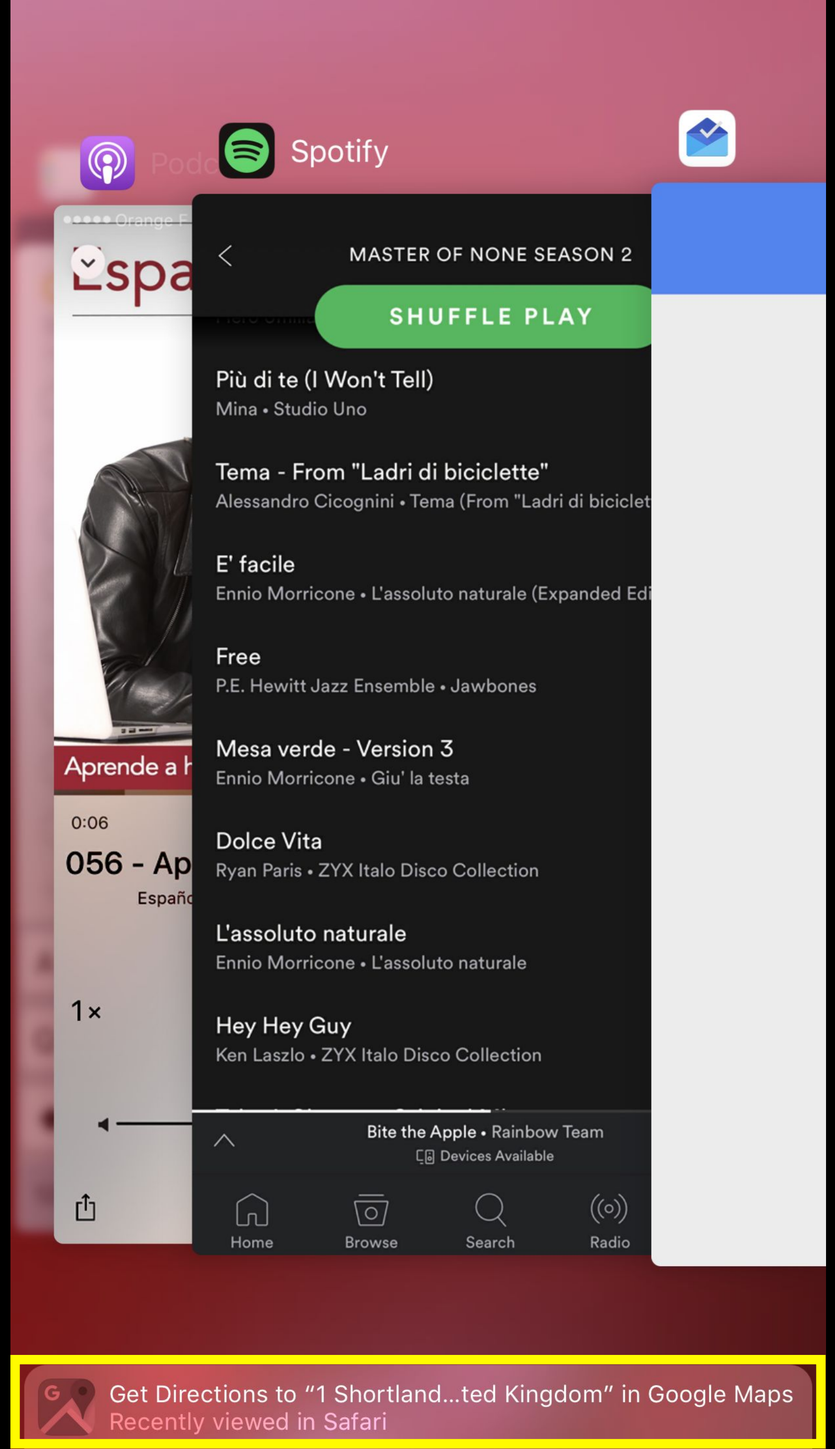
Movies Included with Prime Membership at No Additional Cost See more

CM4 Smartphone Wallet Cases See more

Your Recently Viewed Items and Featured Recommendations

Inspired by your browsing history Page 4 of 10 Start over


●○○○○ Orange F
11:34
Search
SIRI SUGGESTIONS
Show More
Richard
Adeline
Camille
Anne
Skype
Messenger
Hopper
Apple Store
NEARBY
Lunch
Coffee
Shopping
Petrol
NEWS
Show More
Cristiano Ronaldo à Blanc : «J'aimerais beaucoup travailler a...
C'est l'histoire d'un joueur qui veut quitter le Real Madrid. Et depuis mardi


Spotify
MASTER OF NONE SEASON 2
SHUFFLE PLAY
Più di te (I Won't Tell)
Mina • Studio Uno
Tema - From "Ladri di biciclette"
Alessandro Cicognini • Tema (From "Ladri di biciclet
E' facile
Ennio Morricone • L'assoluto naturale (Expanded Edi
Free
P.E. Hewitt Jazz Ensemble • Jawbones
Mesa verde - Version 3
Ennio Morricone • Giu' la testa
Dolce Vita
Ryan Paris • ZYX Italo Disco Collection
L'assoluto naturale
Ennio Morricone • L'assoluto naturale
Hey Hey Guy
Ken Laszlo • ZYX Italo Disco Collection
Bite the Apple • Rainbow Team
Devices Available
Home
Browse
Search
Radio
0:06
1×
Get Directions to “1 Shortland...ted Kingdom” in Google Maps
Recently viewed in Safari

PAR – BUE
Jun 1 – Jun 17

You should keep watching this trip.

You will be notified when deals are available or prices start to rise.

Watching

Price Prediction

Today until Apr 7
Prices will vary over the next 4 months, dropping as low as **671 €** (savings of **70 €**).

Apr 8 – May 26
Prices will probably rise about **86 €** after April 8.

May 27 – Jun 1
Prices are expected to climb at least **438 €** after May 27.

Reminders

Search

Inbox

Snoozed

Done

Drafts

Sent

Reminders

Trash

Spam

Contacts

Trips

PDF E-book Requested

forward the download url to your electronic book's edition & additional code samples referenced within the Amazon.com product detail page

Greg to bml-ebook

Oct 28

Hi, Louis.

Please forward the download url to your electronic book's edition & additional code samples referenced within the Amazon.com product detail page.

Thank you very much.

Greg Newman
gnwmn@yahoo.com

What do developers *actually* need to know about Machine Learning?

Which algorithms do you need to know?

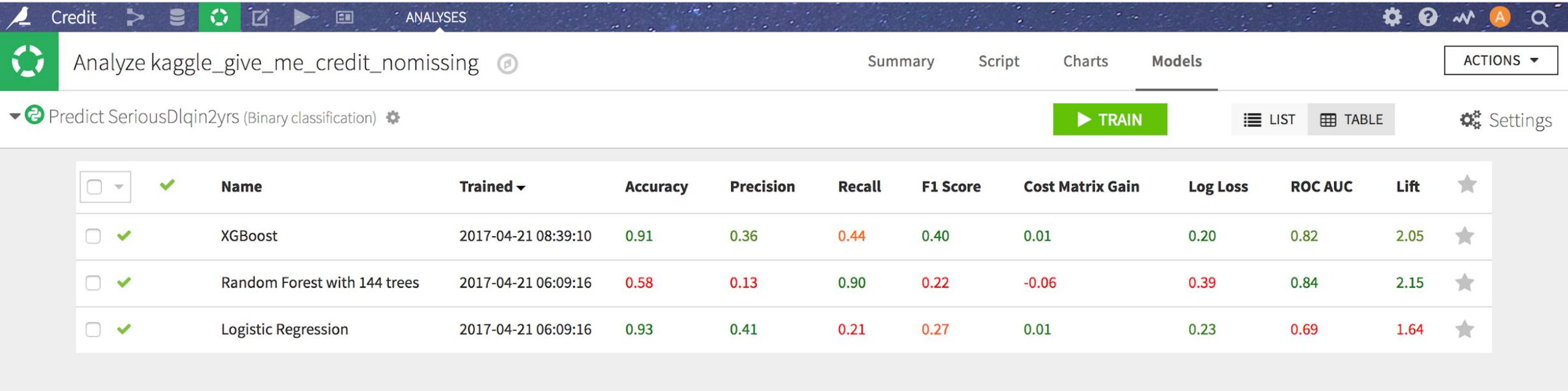
- Do we Need Hundreds of Classifiers to Solve Real World Classification Problems?
- **Random forests** most likely to perform best
- Based on **decision trees**
- Implemented in all ML tools

How much do you need to know?

- Intuitions — what makes learning possible?
- High-level understanding of learning algorithm
- Main parameters
- Possibilities and limitations

how do we quantify that?

Model evaluation



Name	Trained	Accuracy	Precision	Recall	F1 Score	Cost Matrix Gain	Log Loss	ROC AUC	Lift
XGBoost	2017-04-21 08:39:10	0.91	0.36	0.44	0.40	0.01	0.20	0.82	2.05
Random Forest with 144 trees	2017-04-21 06:09:16	0.58	0.13	0.90	0.22	-0.06	0.39	0.84	2.15
Logistic Regression	2017-04-21 06:09:16	0.93	0.41	0.21	0.27	0.01	0.23	0.69	1.64