goto; chicago

Machine Learning

without a PhD

follow us @gotochgo

Machine Learning without a PhD

10:15-11:00
Relating to Machine Learning
Stefan Veis Pennerup

11:15-12:00
Exploring StackOverflow Data
Evelina Gabasova

13:00-13:45
Developing a ML model
Kevin Tsai

14:00-14:45
Life and Death Decisions: Testing Data Science
Phil Winder

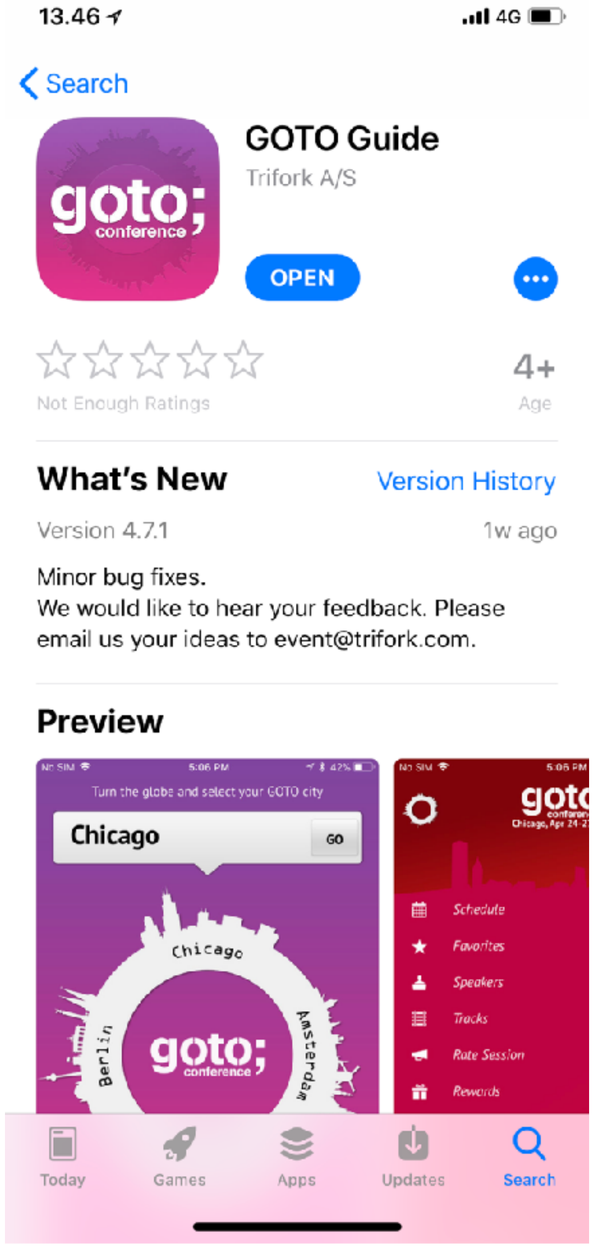
15:15-16:00
Production Model Deployment
Juliet Hougland

16:15-16:45
Delivering AI on Code: Live Demo of source{d}
Francesc Campoy

Housekeeping rules

- Sessions are 45 minutes, including questions
- You can ask questions through the GOTO app
- The track host will read questions at the end of the presentation

GOTO Guide



Download on the App Store

GET IT ON Google Play

GOTO Guide
Trifork
PEGI 3

UNINSTALL OPEN

5 THOUSAND Downloads
3.8 35
Business
Similar

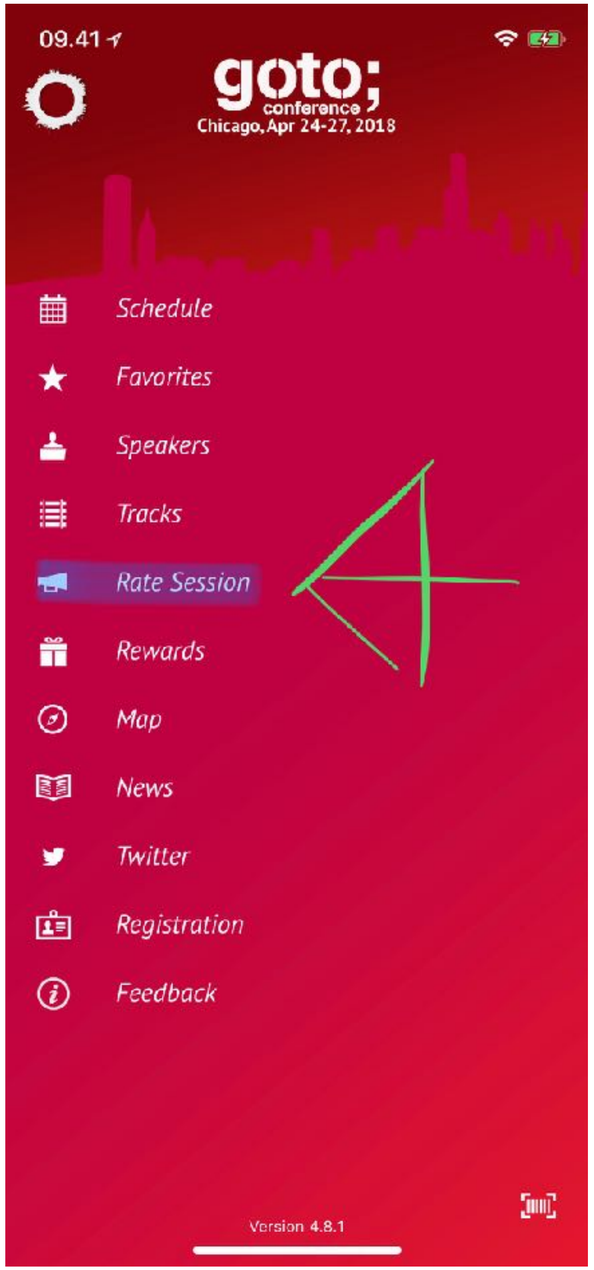
The official Android app for GOTO Conferences

WHAT'S NEW
bug fixes.

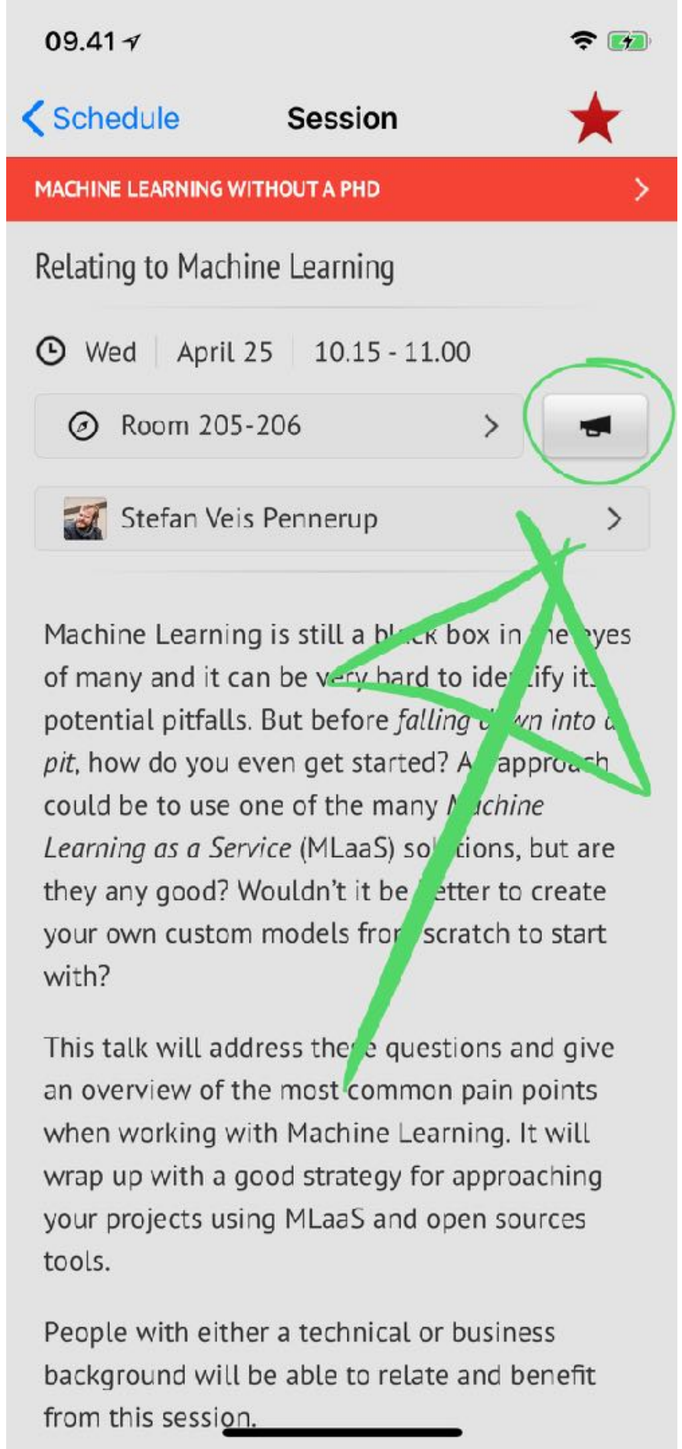
GOTO Guide

Select Chicago on the Wheel

Find “Rate Session” in the menu



Click to ask question and rate the session



goto; chicago

Relating to Machine Learning

Stefan Veis Pennerup

follow us @gotochgo

Who is this talk for?

TRIFORK.
...think software

What is Machine Learning?

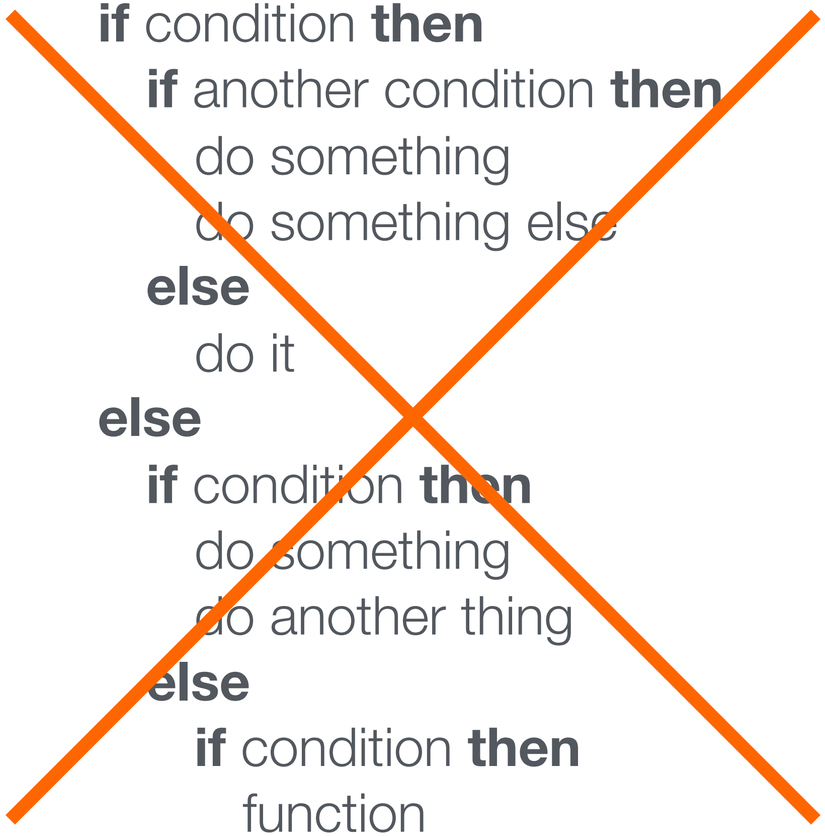
TRIFORK.
...think software

Rule based

```
if condition then
   if another condition then
      do something
      do something else
   else
      do it
else
   if condition then
      do something
      do another thing
   else
      if condition then
         function
```

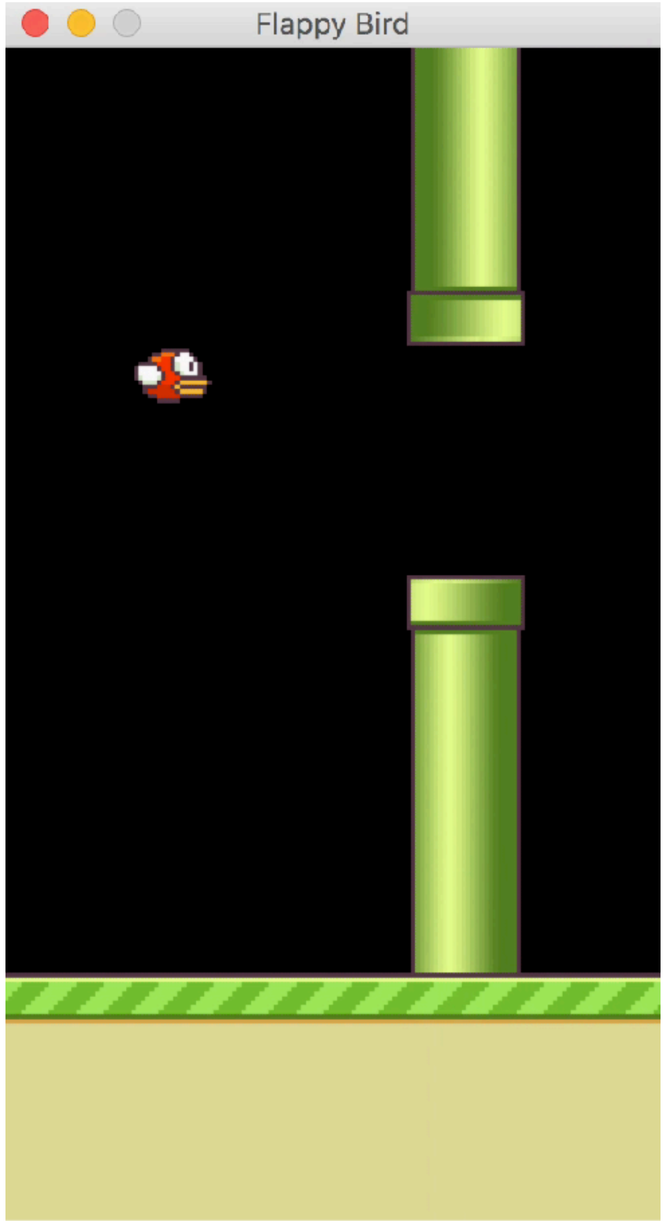
TRIFORK.
...think software

Rule based



TRIFORK.
...think software

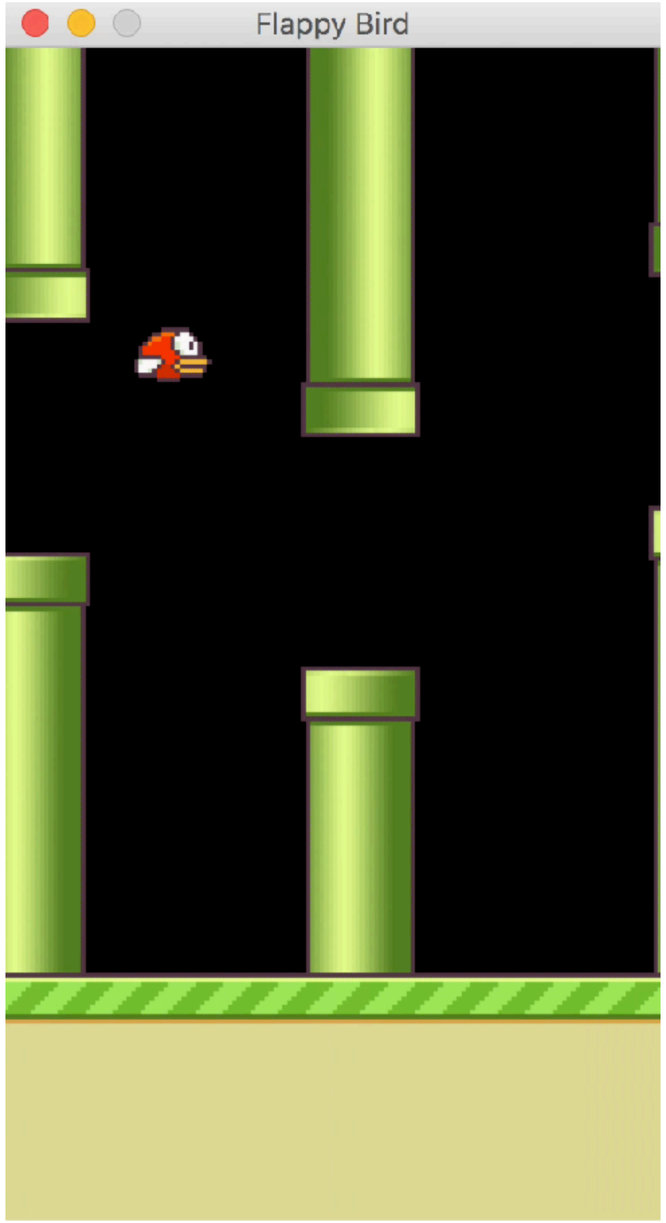
Learn by example



https://github.com/yenchenlin/DeepLearningFlappyBird

TRIFORK.
...think software

Learn by example



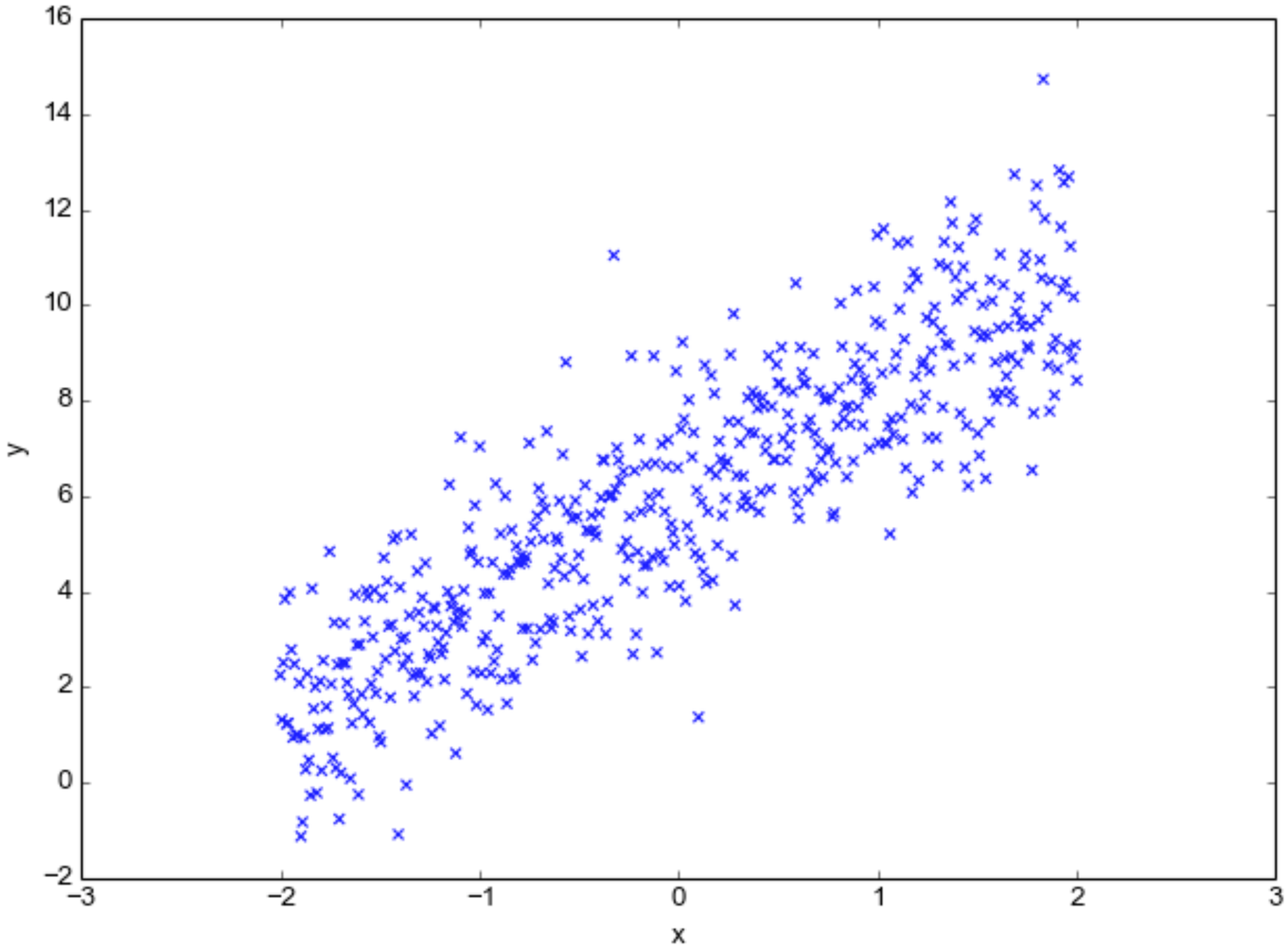
https://github.com/yenchenlin/DeepLearningFlappyBird

TRIFORK.
...think software

Linear Regression

TRIFORK.
...think software

The goal of linear regression



https://eli.thegreenplace.net/2016/linear-regression/

TRIFORK.
...think software

Linear regression without ML

$$a = \frac{\sum_{i=1}^{n}(x_i - \overline{x}) * (y_i - \overline{y})}{\sum_{i=1}^{n}(x_i - \overline{x})^2}$$

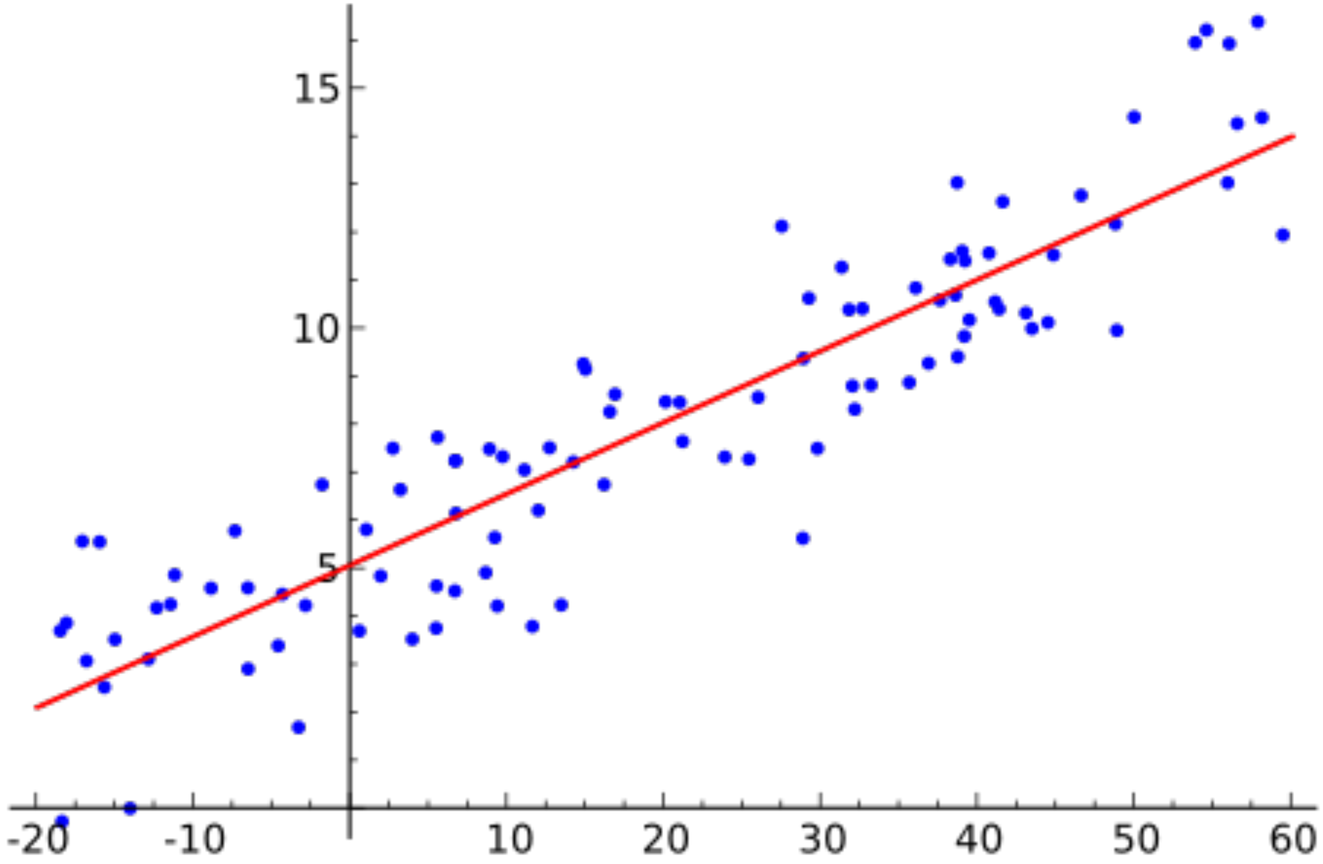
$$b = \overline{y} - a\overline{x}$$

TRIFORK.
...think software

Linear regression without ML

$$a = \frac{\sum_{i=1}^{n}(x_i - \overline{x}) * (y_i - \overline{y})}{\sum_{i=1}^{n}(x_i - \overline{x})^2}$$

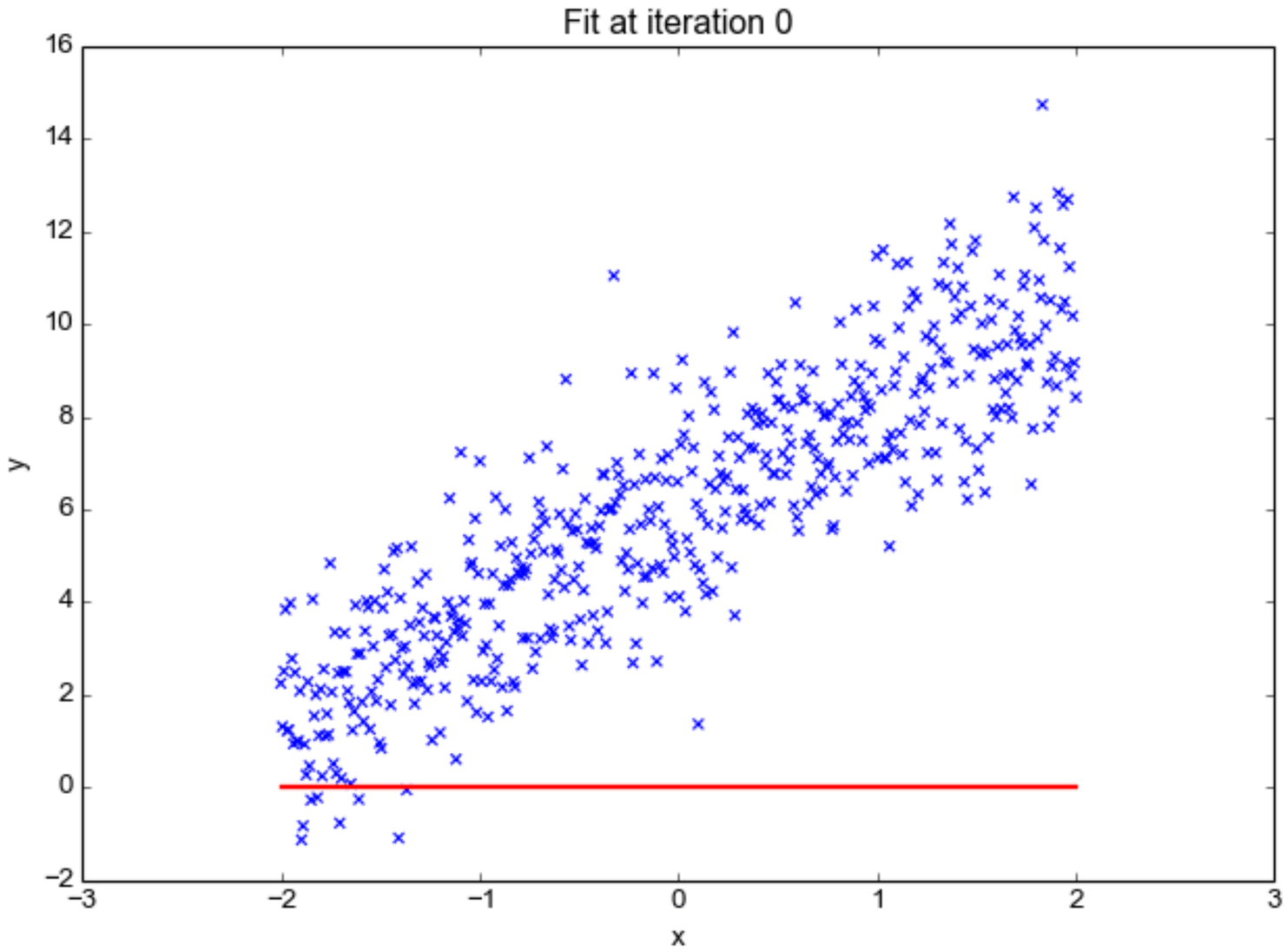
$$b = \overline{y} - a\overline{x}$$



https://en.wikipedia.org/wiki/Linear_regression#/media/File:Linear_regression.svg

TRIFORK.
...think software

Linear regression with ML



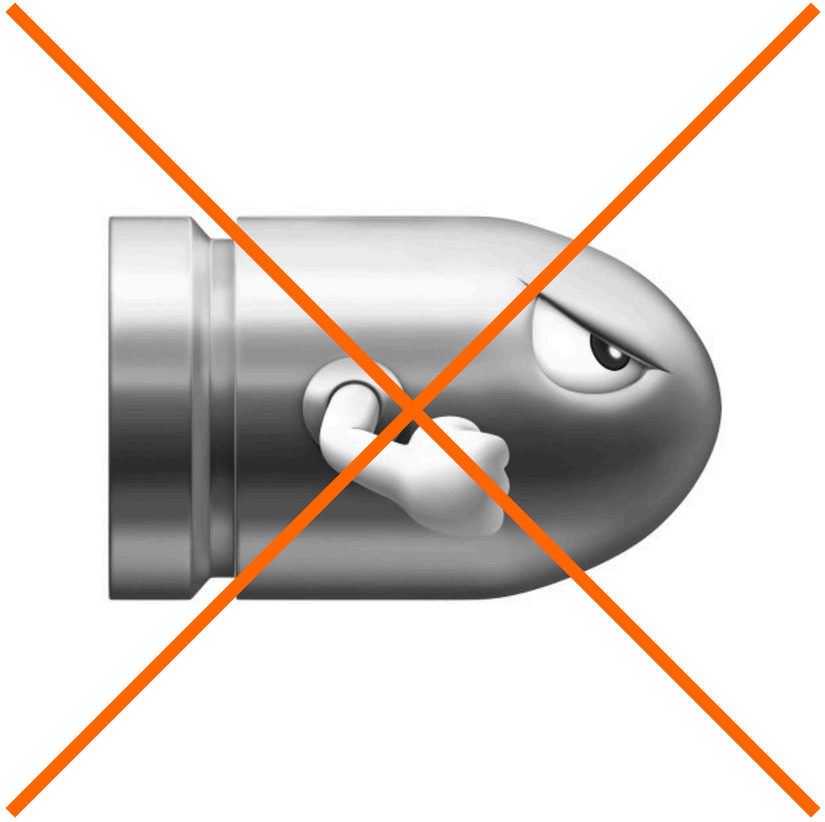
https://eli.thegreenplace.net/2016/linear-regression/

TRIFORK.
...think software

Machine Learning intuitions

TRIFORK.
...think software

Machine Learning is not a silver bullet 🔫

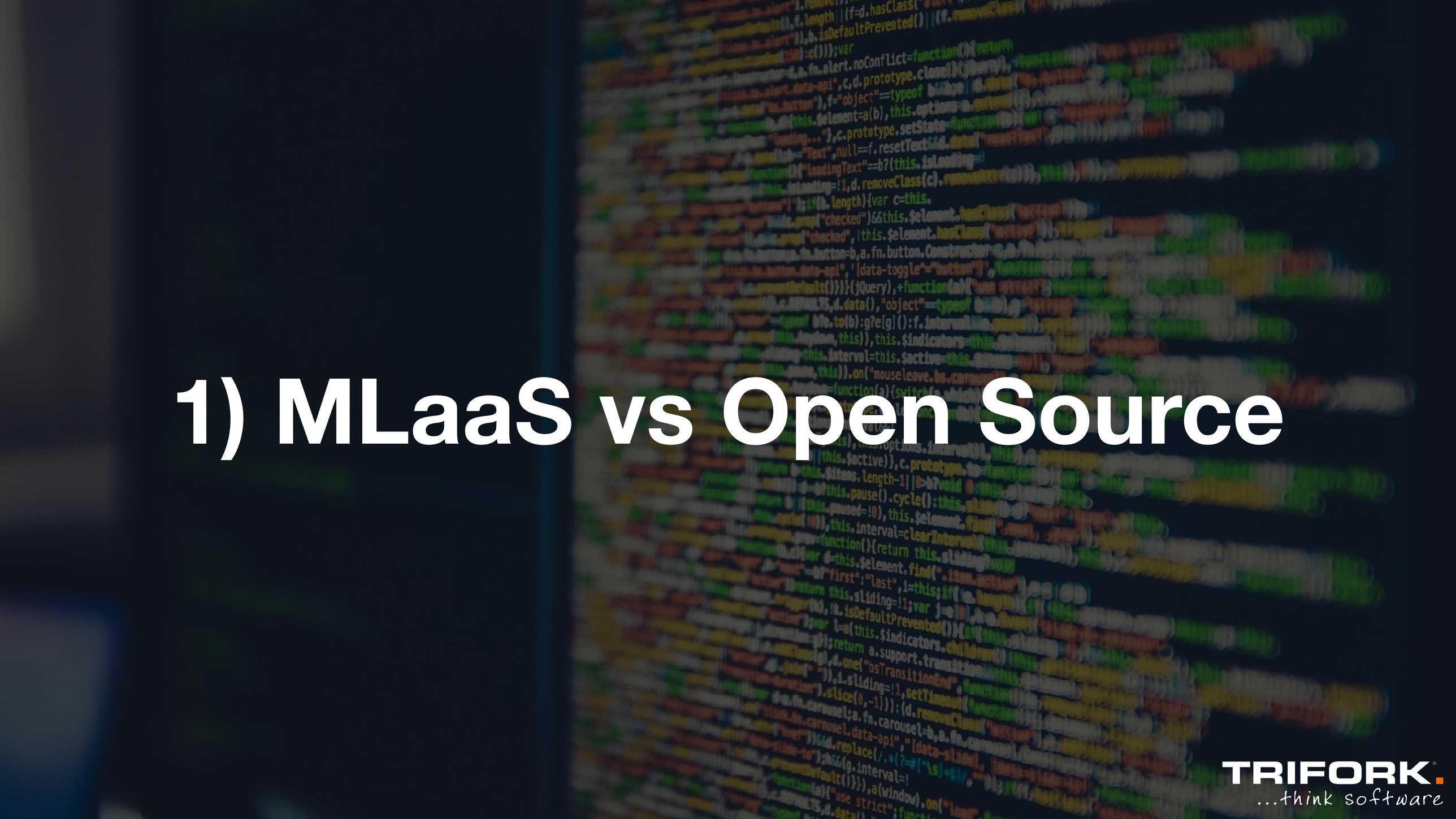


ML is just another tool in the toolbox 🛠

5 intuitions

1. MLaaS vs Open Source
2. Image recognition of complex objects
3. Vulnerabilities
4. Imbalanced datasets
5. Human bias

1) MLaaS vs Open Source



TRIFORK.
...think software

MLaaS vs Open Source

Machine Learning as a Service (MLaaS)

- Out-of-the-box solutions
- Pay per usage

TRIFORK.
...think software

MLaaS vs Open Source

Machine Learning as a Service (MLaaS)

- Out-of-the-box solutions
- Pay per usage

Open Source

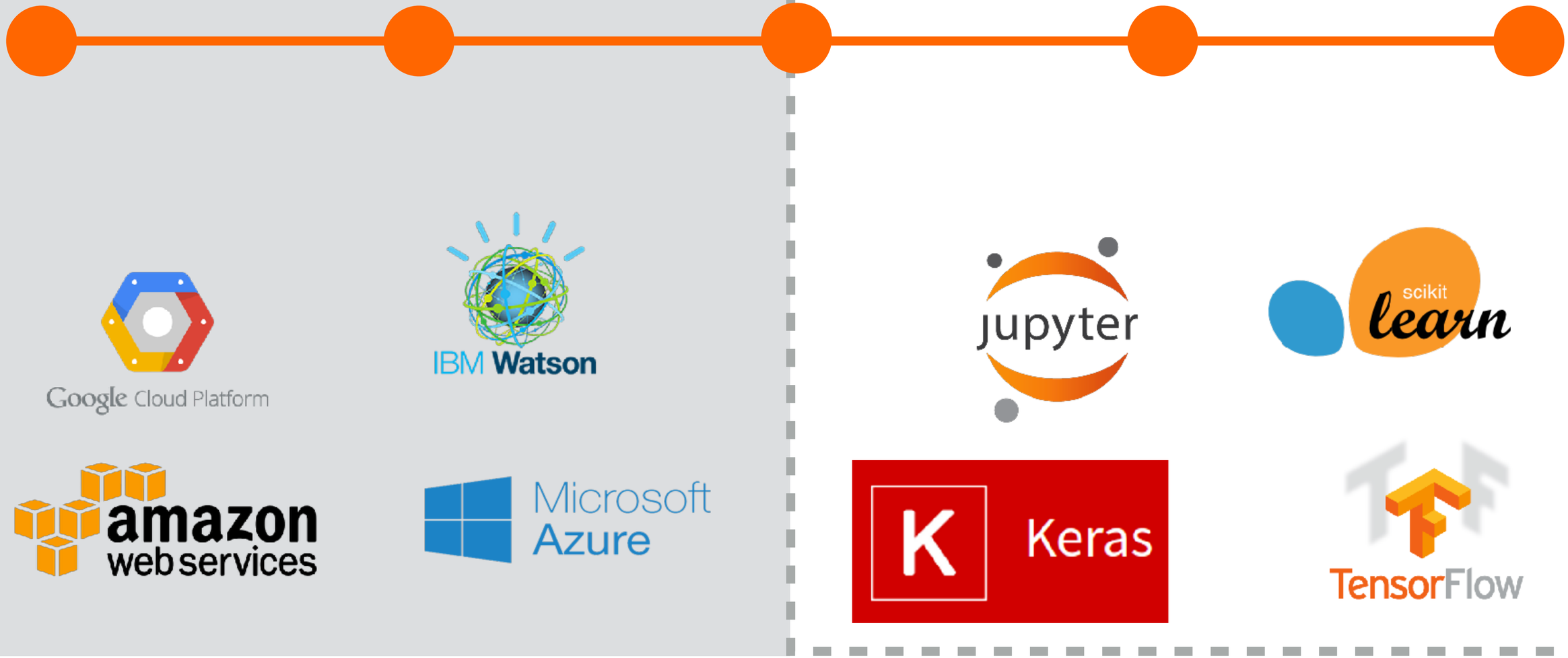
- Custom models for your use case
- Requires the necessary skills

TRIFORK.
...think software

Experiment with the whole scale

Machine Learning as a Service (MLaaS)

Open Source



Experiment with the whole scale

Machine Learning as a Service (MLaaS)

Open Source

Ready to use APIs

TRIFORK.
...think software

Experiment with the whole scale

Machine Learning as a Service (MLaaS)

Open Source

Ready to use APIs

Custom training with MLaaS

TRIFORK.
...think software

Experiment with the whole scale

Machine Learning as a Service (MLaaS)

Open Source

Ready to use APIs

Custom training with MLaaS

Pre-trained models

Google Cloud Platform

IBM Watson

jupyter

scikit learn

amazon web services

Microsoft Azure

Keras

TensorFlow

TRIFORK
...think software

Experiment with the whole scale

Machine Learning as a Service (MLaaS)

Open Source

Ready to use APIs

Custom training with MLaaS

Pre-trained models

Transfer Learning

Google Cloud Platform

IBM Watson

jupyter

scikit learn

amazon web services

Microsoft Azure

Keras

TensorFlow

TRIFORK.
...think software

Experiment with the whole scale

Machine Learning as a Service (MLaaS)

Open Source

Ready to use APIs

Custom training with MLaaS

Pre-trained models

Transfer Learning

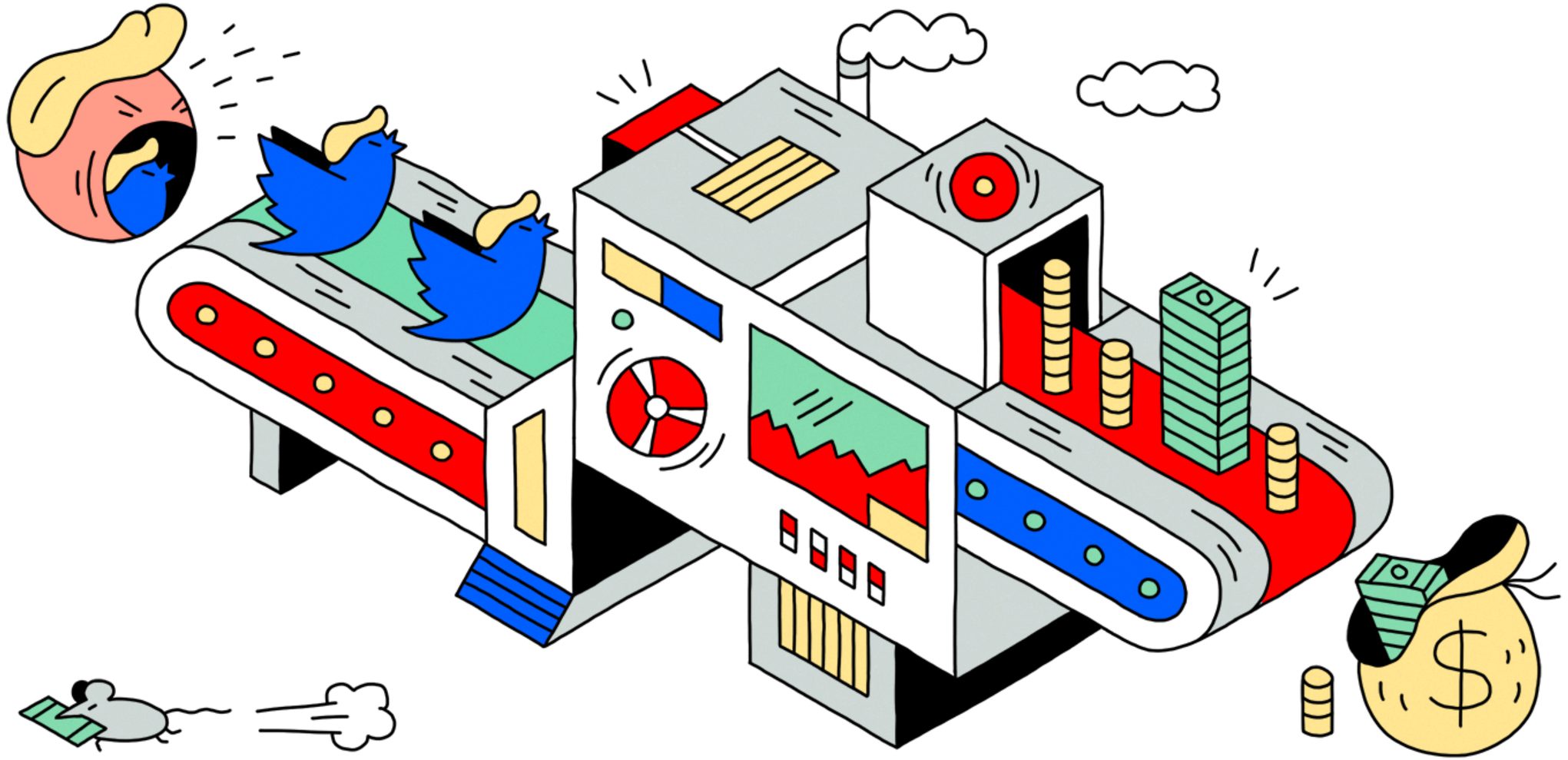
Custom model

TRIFORK.
...think software

What’s feasible with MLaaS?

TRIFORK.
...think software

Trump2Cash



https://github.com/maxbbraun/trump2cash

TRIFORK.
...think software

2) Recognising complex objects

Banana

Apple

Banana

https://gist.github.com/yrevar/942d3a0ac09ec9e5eb3a

```
947: 'mushroom',
948: 'Granny Smith',
949: 'strawberry',
950: 'orange',
951: 'lemon',
952: 'fig',
953: 'pineapple, ananas',
954: 'banana',
955: 'jackfruit, jak, jack',
956: 'custard apple',
957: 'pomegranate',
958: 'hay',
959: 'carbonara',
960: 'chocolate sauce, chocolate
961: 'dough',
962: 'meat loaf, meatloaf',
```

Apple - Granny Smith

Norfolk terrier

https://www.akc.org/dog-breeds/norfolk-terrier/

TRIFORK.
...think software

Norwich Terrier

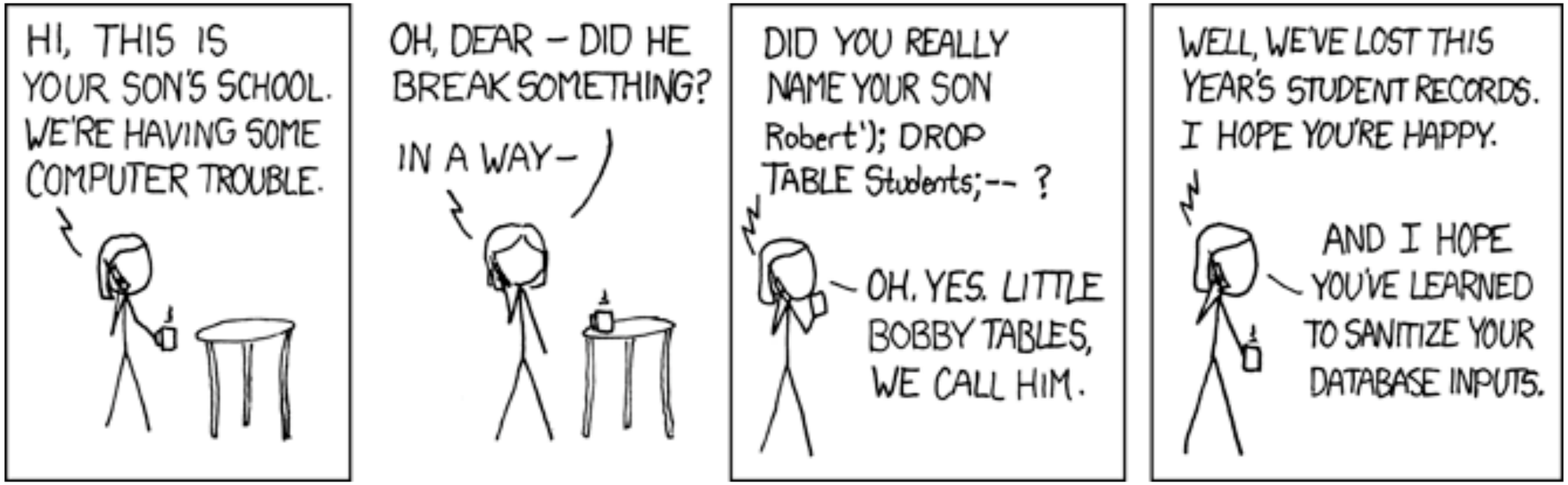
http://www.akc.org/dog-breeds/norwich-terrier/

TRIFORK.
...think software

3) Vulnerabilities

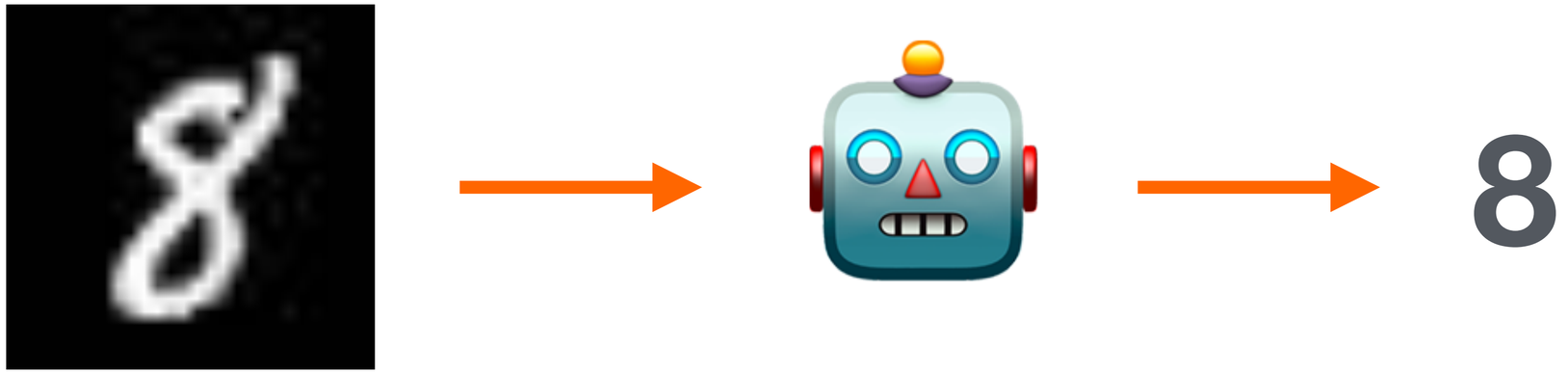
TRIFORK.
...think software

Previous IT vulnerabilities



https://xkcd.com/327/

Neural Networks can easily be fooled



A. Nguyen, J. Yosinski, J. Clune. Deep Neural Networks are Easily Fooled: High Confidence Predictions for Unrecognizable Images. 2015.

Neural Networks can easily be fooled

A. Nguyen, J. Yosinski, J. Clune. Deep Neural Networks are Easily Fooled: High Confidence Predictions for Unrecognizable Images. 2015.

Neural Networks can easily be fooled

Gibberish

A. Nguyen, J. Yosinski, J. Clune. Deep Neural Networks are Easily Fooled: High Confidence Predictions for Unrecognizable Images. 2015.

TRIFORK.
...think software

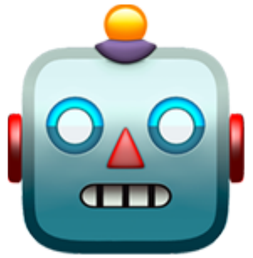
Neural Networks can easily be fooled

Gibberish

8

A. Nguyen, J. Yosinski, J. Clune. Deep Neural Networks are Easily Fooled: High Confidence Predictions for Unrecognizable Images. 2015.

Neural Networks can easily be fooled



Gibberish

8

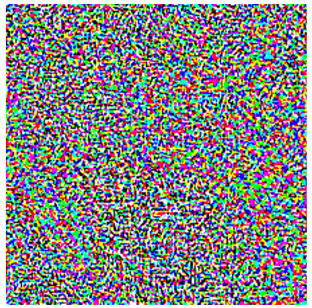
99.99% confidence

A. Nguyen, J. Yosinski, J. Clune. Deep Neural Networks are Easily Fooled: High Confidence Predictions for Unrecognizable Images. 2015.

TRIFORK. ...think software

Adversarial attack

+



Panda
57.7% confidence

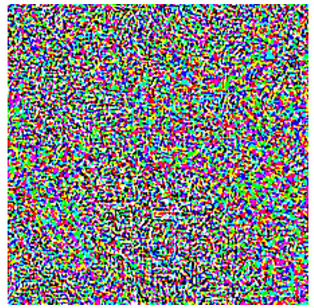
I. J. Goodfellow, J. Shlens, C. Szegedy. Explaining and Harnessing Adversarial Examples. 2015.

TRIFORK.
...think software

Adversarial attack

Panda
57.7% confidence

\+ 0.7% of



=

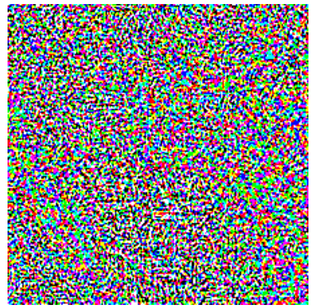
I. J. Goodfellow, J. Shlens, C. Szegedy. Explaining and Harnessing Adversarial Examples. 2015.

TRIFORK.
...think software

Adversarial attack

Panda
57.7% confidence

\+ 0.7% of



=

Gibbon
99.3% confidence

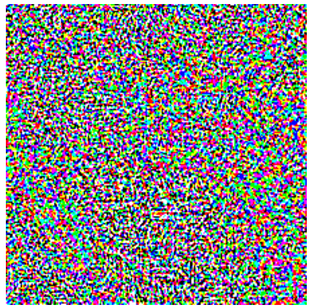
I. J. Goodfellow, J. Shlens, C. Szegedy. Explaining and Harnessing Adversarial Examples. 2015.

TRIFORK.
...think software

Adversarial attack

Panda
57.7% confidence

\+ 0.7% of



=

Gibbon
99.3% confidence

https://en.wikipedia.org/wiki/Gibbon

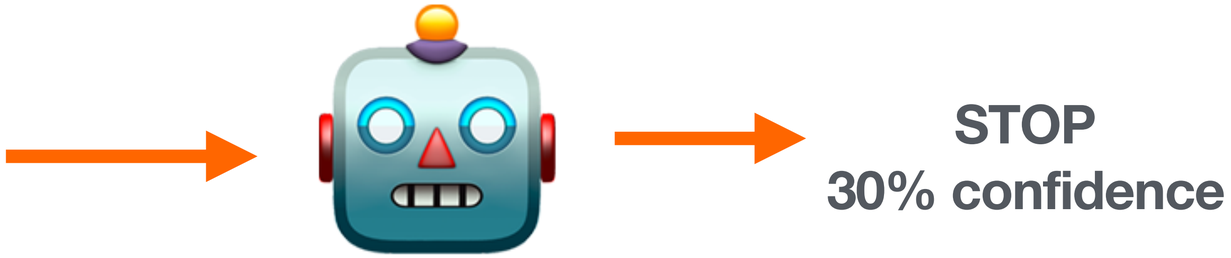
TRIFORK.
...think software

Potential impact of adversarial attack

K. Eykholt, I. Evtimov, E. Fernandes, B. Li, A. Rahmati, C. Xiao, A. Prakash, T. Kohno, D. Song. Robust Physical-World Attacks on Deep Learning Models. 2018.

TRIFORK.
...think software

Potential impact of adversarial attack



STOP
30% confidence

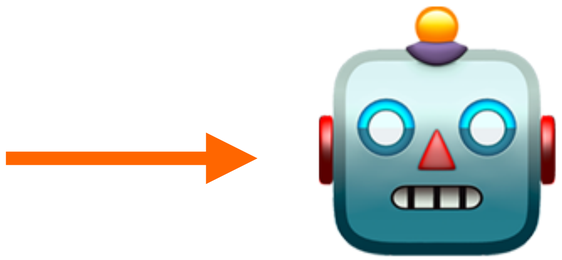
K. Eykholt, I. Evtimov, E. Fernandes, B. Li, A. Rahmati, C. Xiao, A. Prakash, T. Kohno, D. Song. Robust Physical-World Attacks on Deep Learning Models. 2018.

Potential impact of adversarial attack

K. Eykholt, I. Evtimov, E. Fernandes, B. Li, A. Rahmati, C. Xiao, A. Prakash, T. Kohno, D. Song. Robust Physical-World Attacks on Deep Learning Models. 2018.

TRIFORK.
...think software

Potential impact of adversarial attack



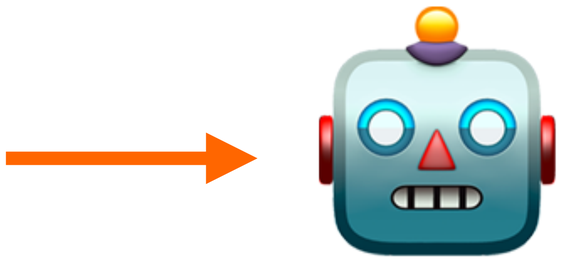
STOP
30% confidence

Speed Limit 45 mph

K. Eykholt, I. Evtimov, E. Fernandes, B. Li, A. Rahmati, C. Xiao, A. Prakash, T. Kohno, D. Song. Robust Physical-World Attacks on Deep Learning Models. 2018.

TRIFORK.
...think software

Potential impact of adversarial attack



STOP
30% confidence

Speed Limit 45 mph
39% confidence

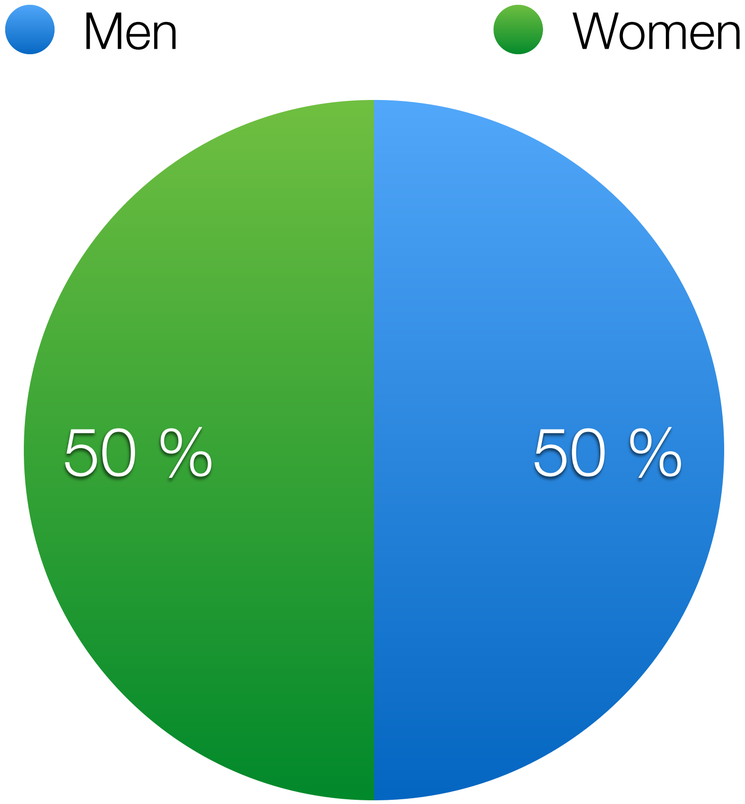
K. Eykholt, I. Evtimov, E. Fernandes, B. Li, A. Rahmati, C. Xiao, A. Prakash, T. Kohno, D. Song. Robust Physical-World Attacks on Deep Learning Models. 2018.

TRIFORK.
...think software

4) Imbalanced datasets

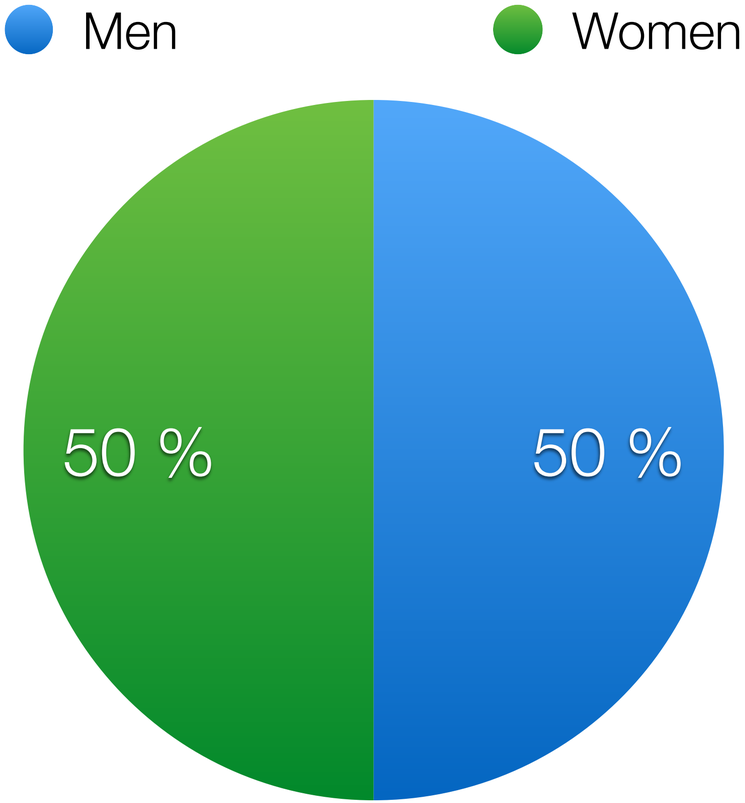
TRIFORK.
...think software

Imbalanced dataset

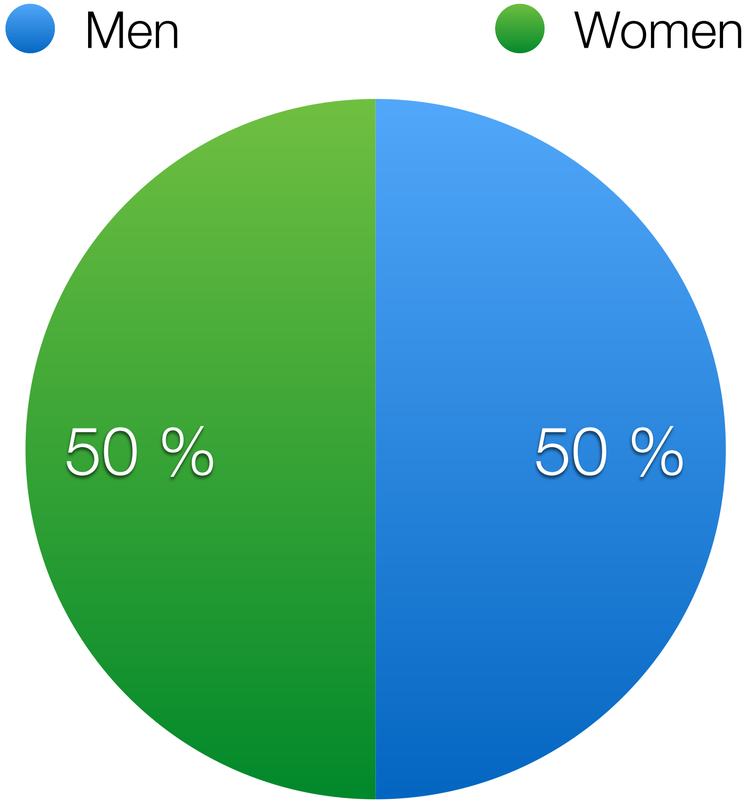


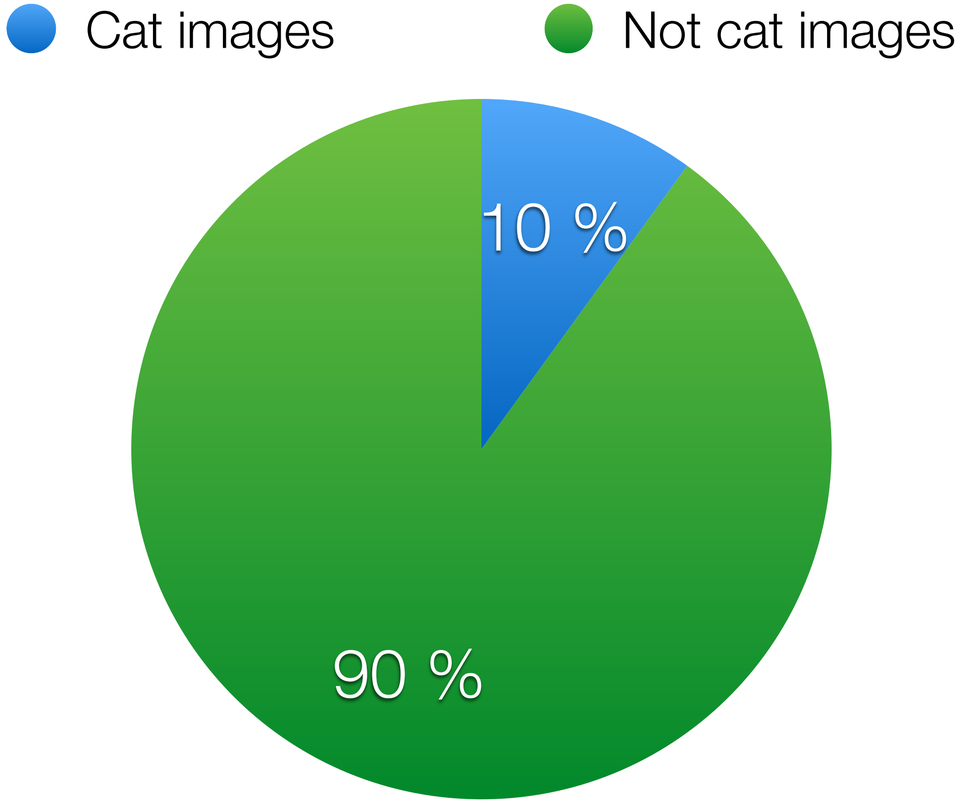
TRIFORK.
...think software

Imbalanced dataset

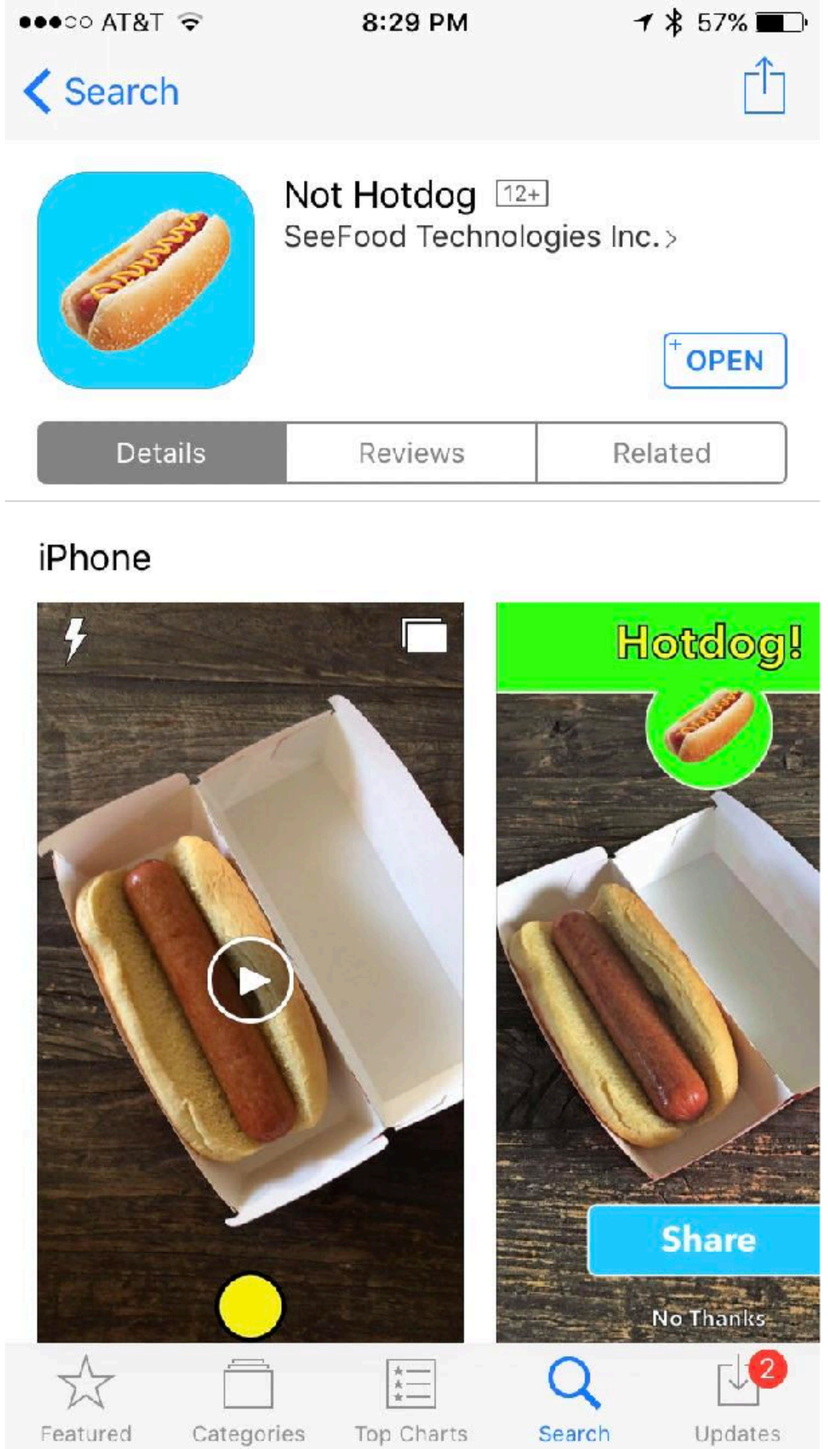


Imbalanced dataset



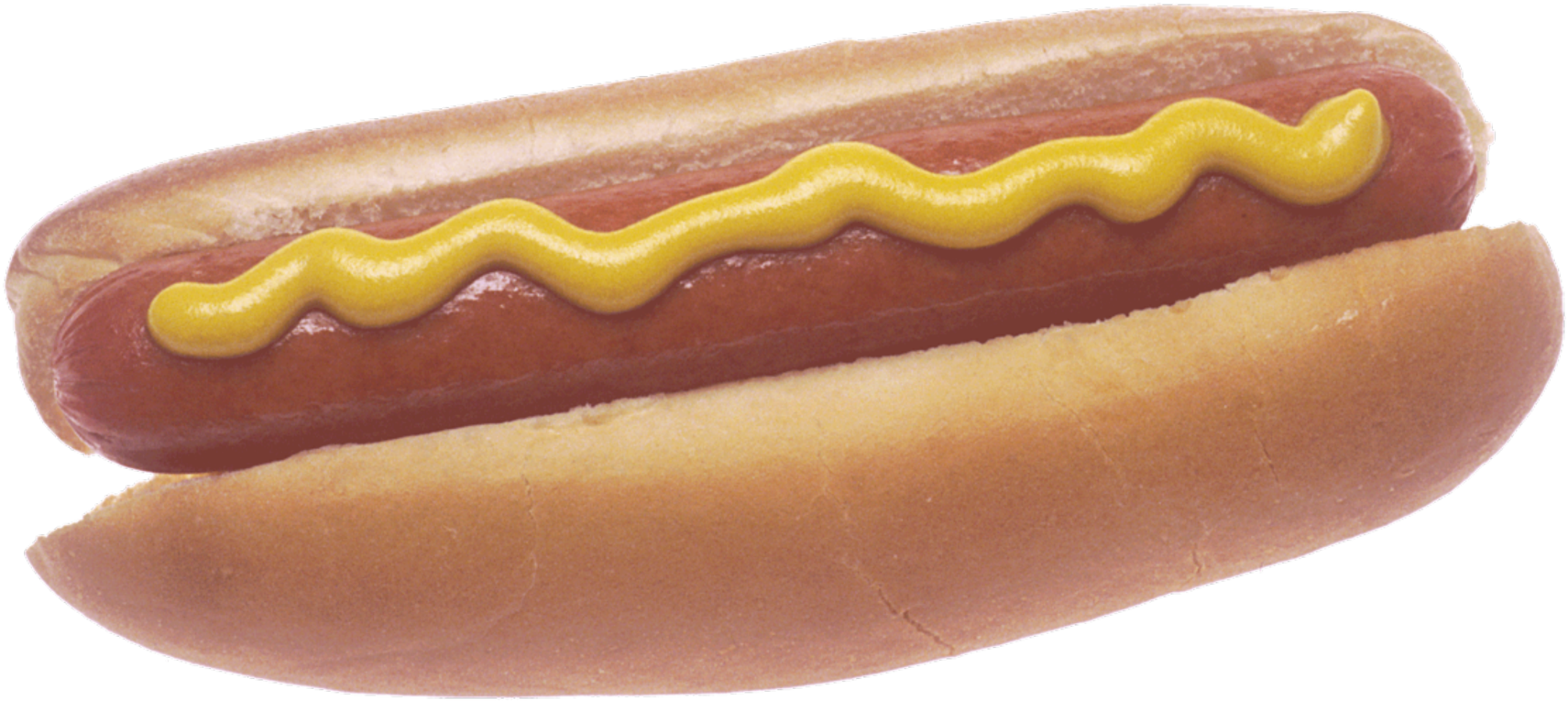


TRIFORK.
...think software



https://itunes.apple.com/app/not-hotdog/id1212457521

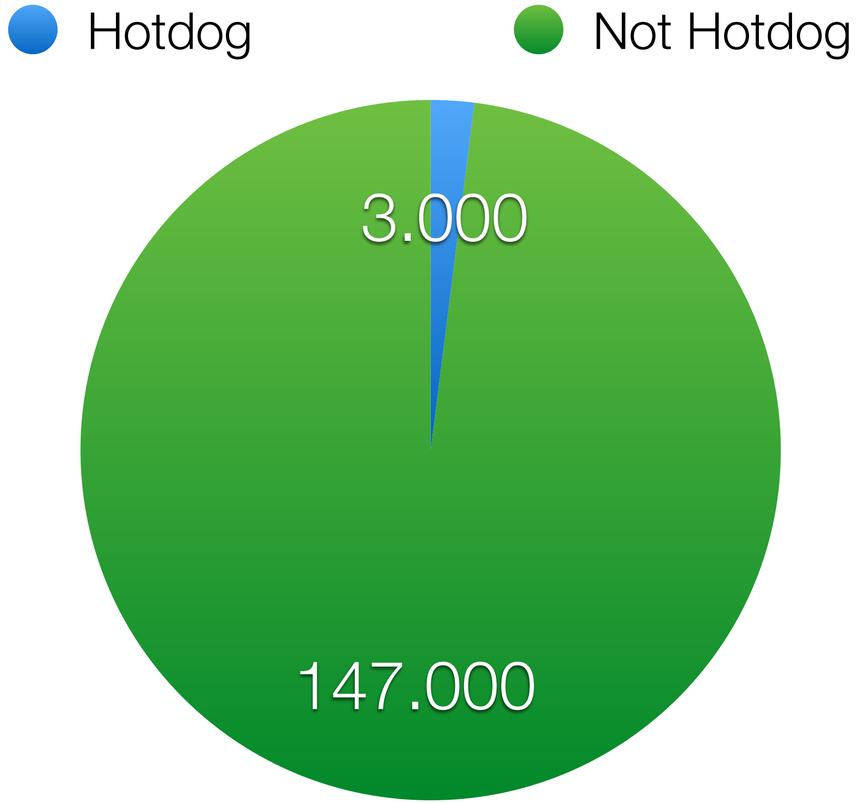
TRIFORK
...think software

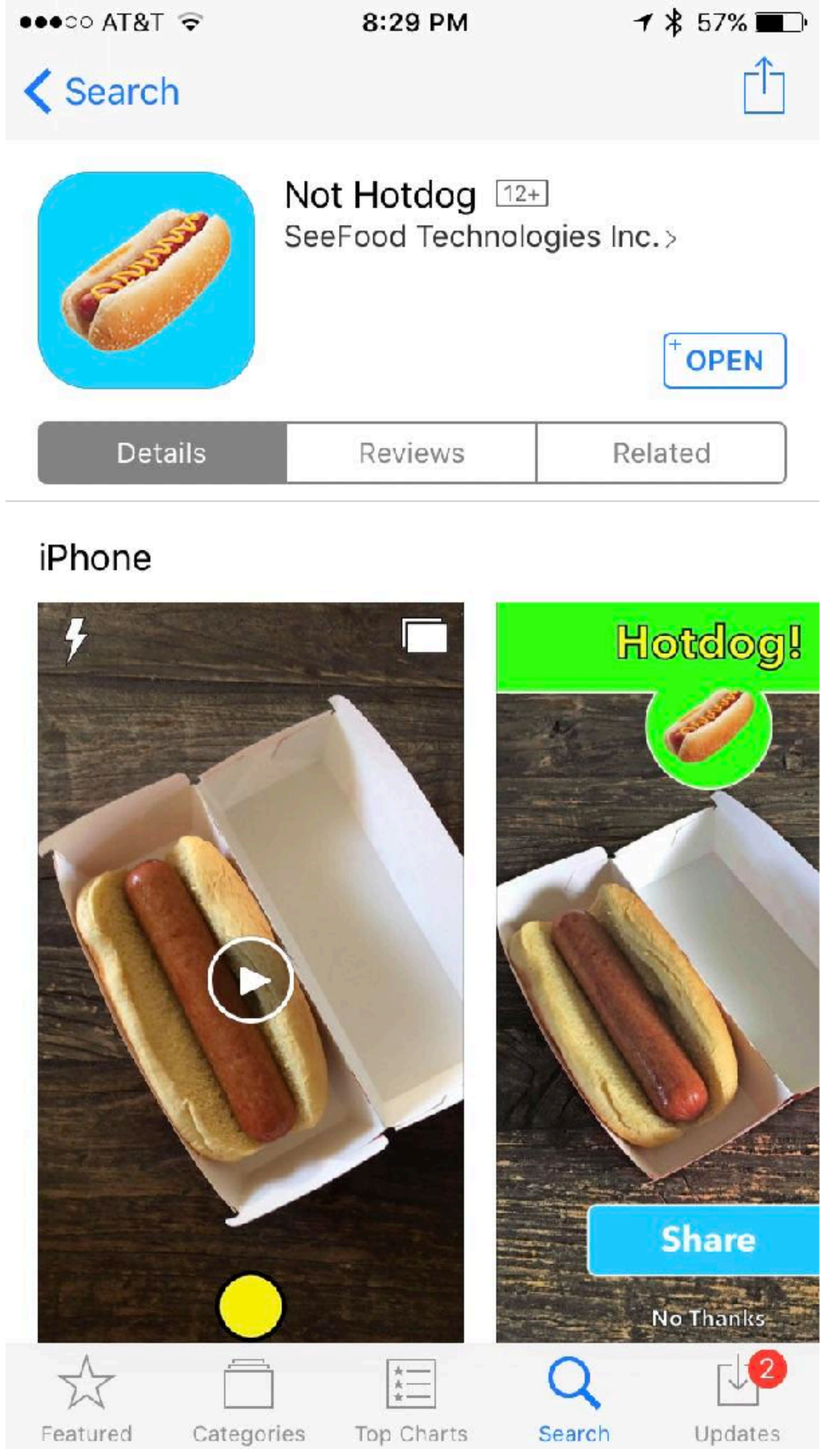


https://en.wikipedia.org/wiki/Hot_dog

TRIFORK.
...think software

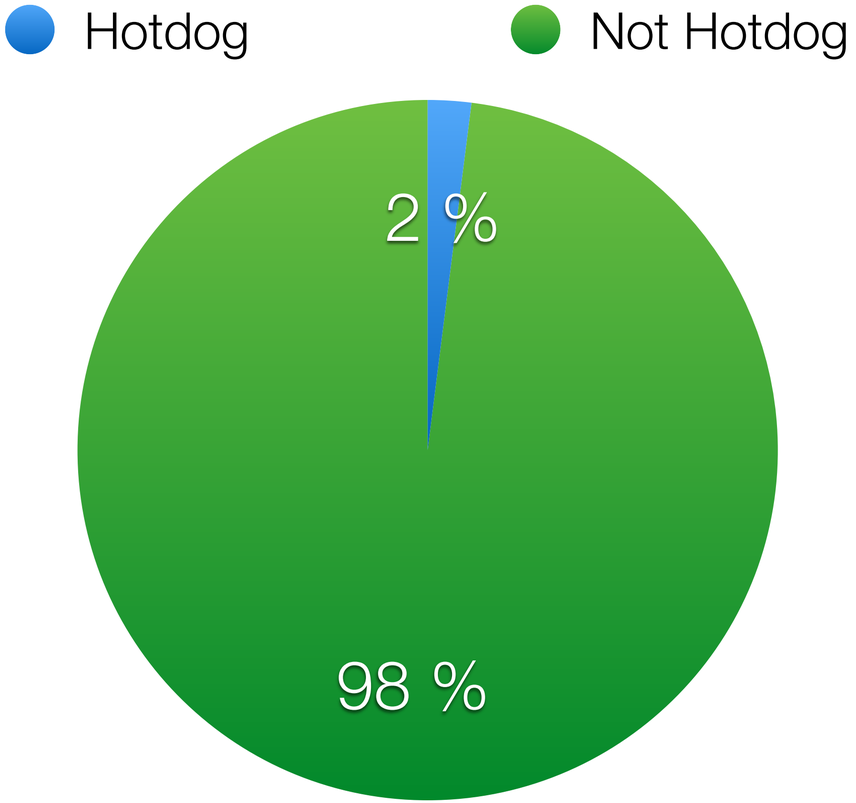
Not Hotdog

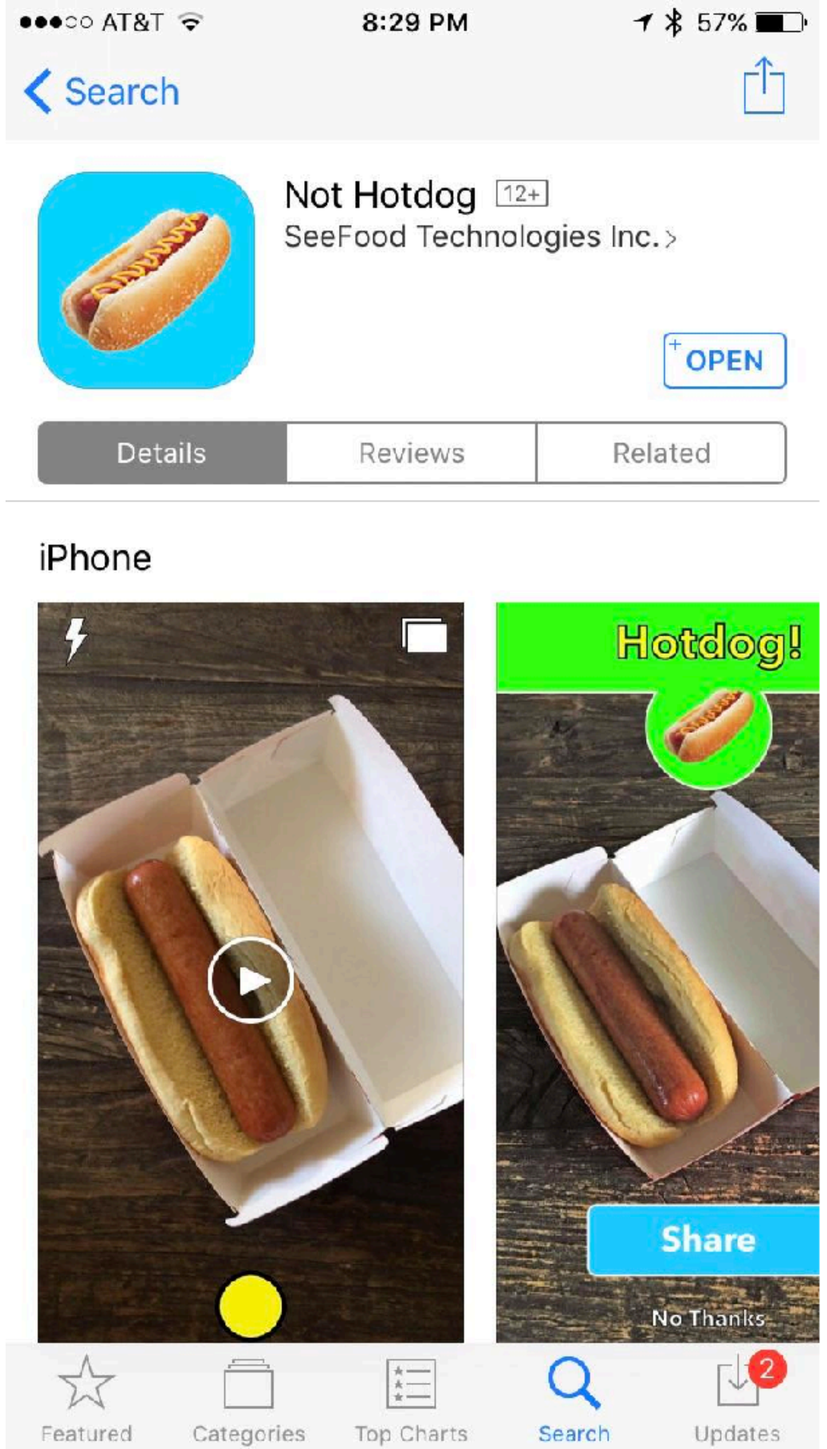




https://medium.com/@timanglade/how-hbos-silicon-valley-built-not-hotdog-with-mobile-tensorflow-keras-react-native-ef03260747f3

Not Hotdog

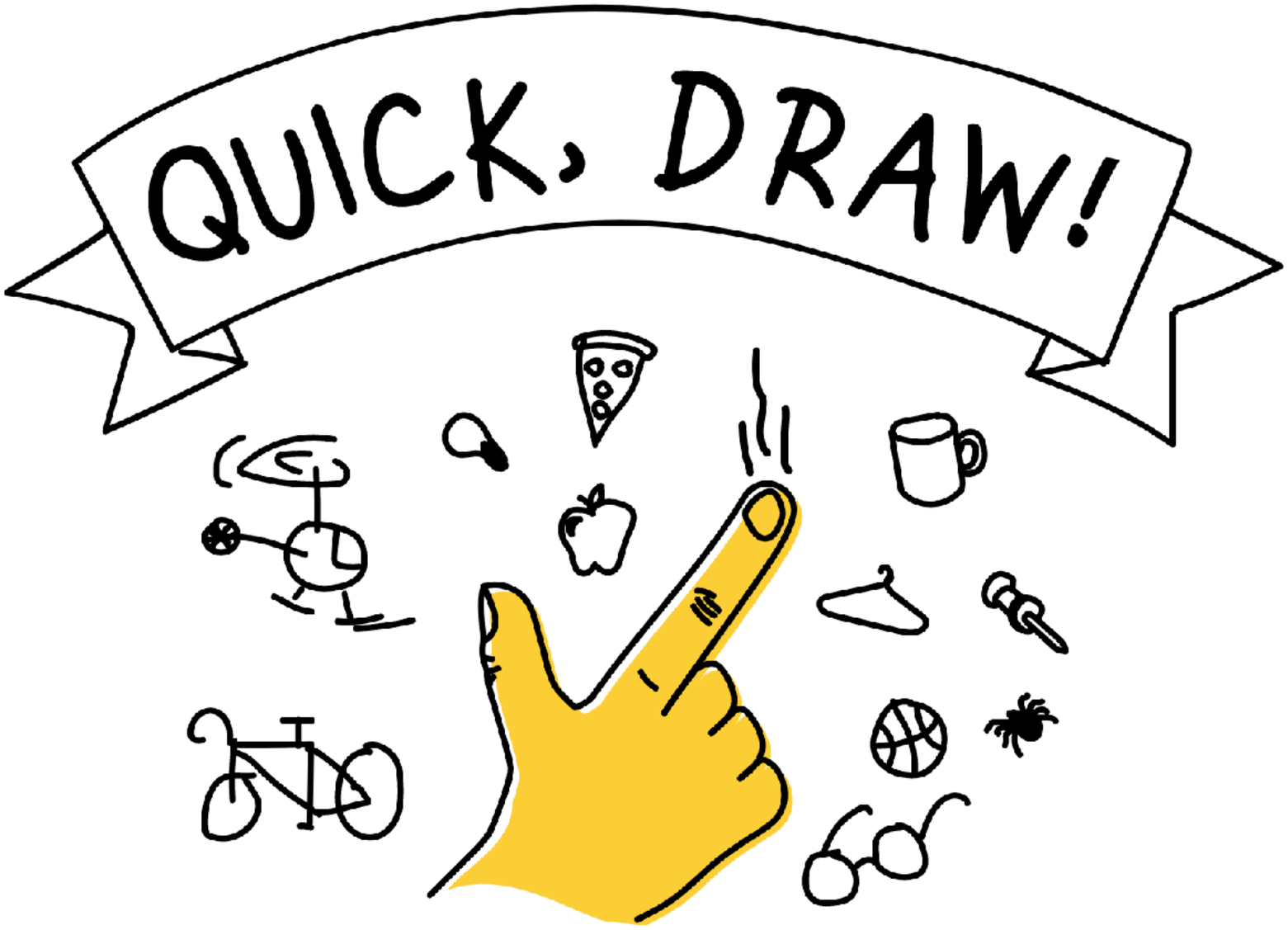




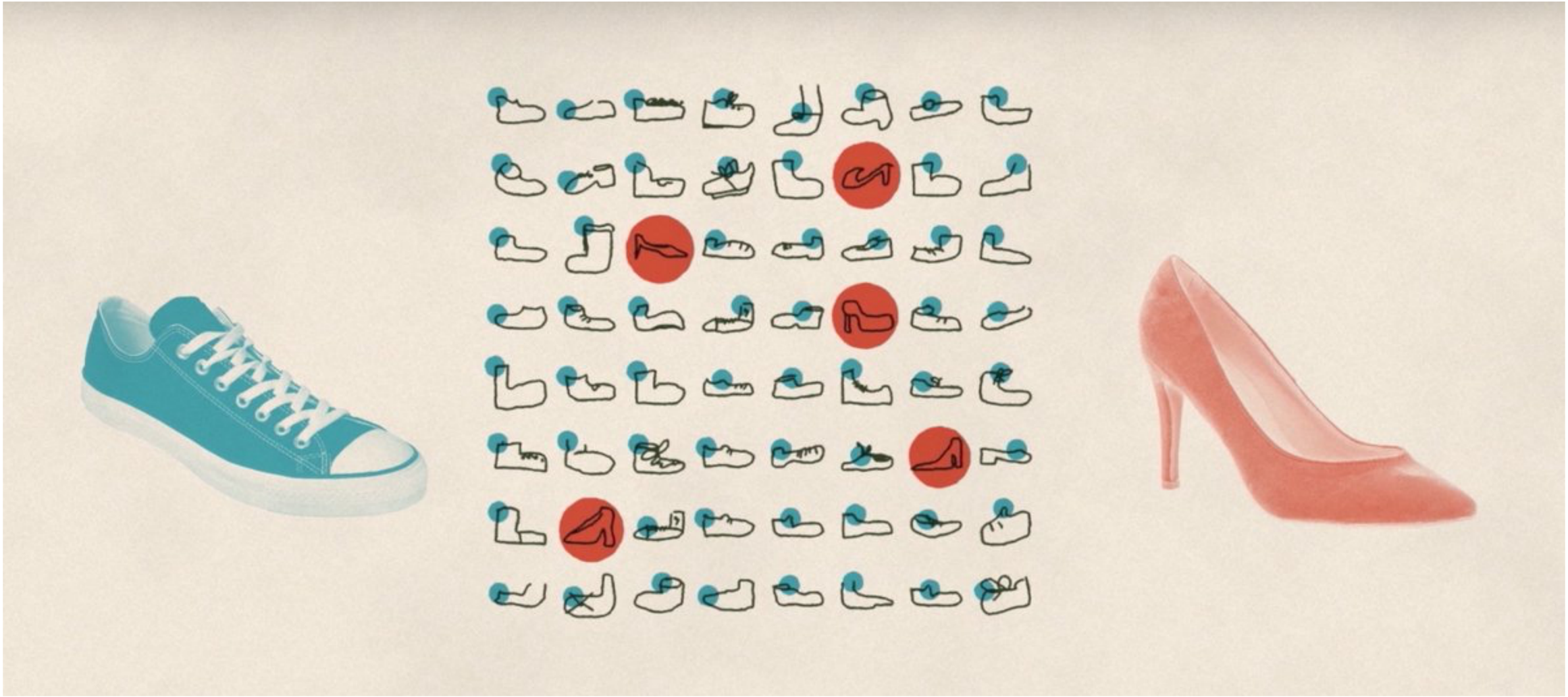
https://medium.com/@timanglade/how-hbos-silicon-valley-built-not-hotdog-with-mobile-tensorflow-keras-react-native-ef03260747f3

5) Human bias

TRIFORK.
...think software

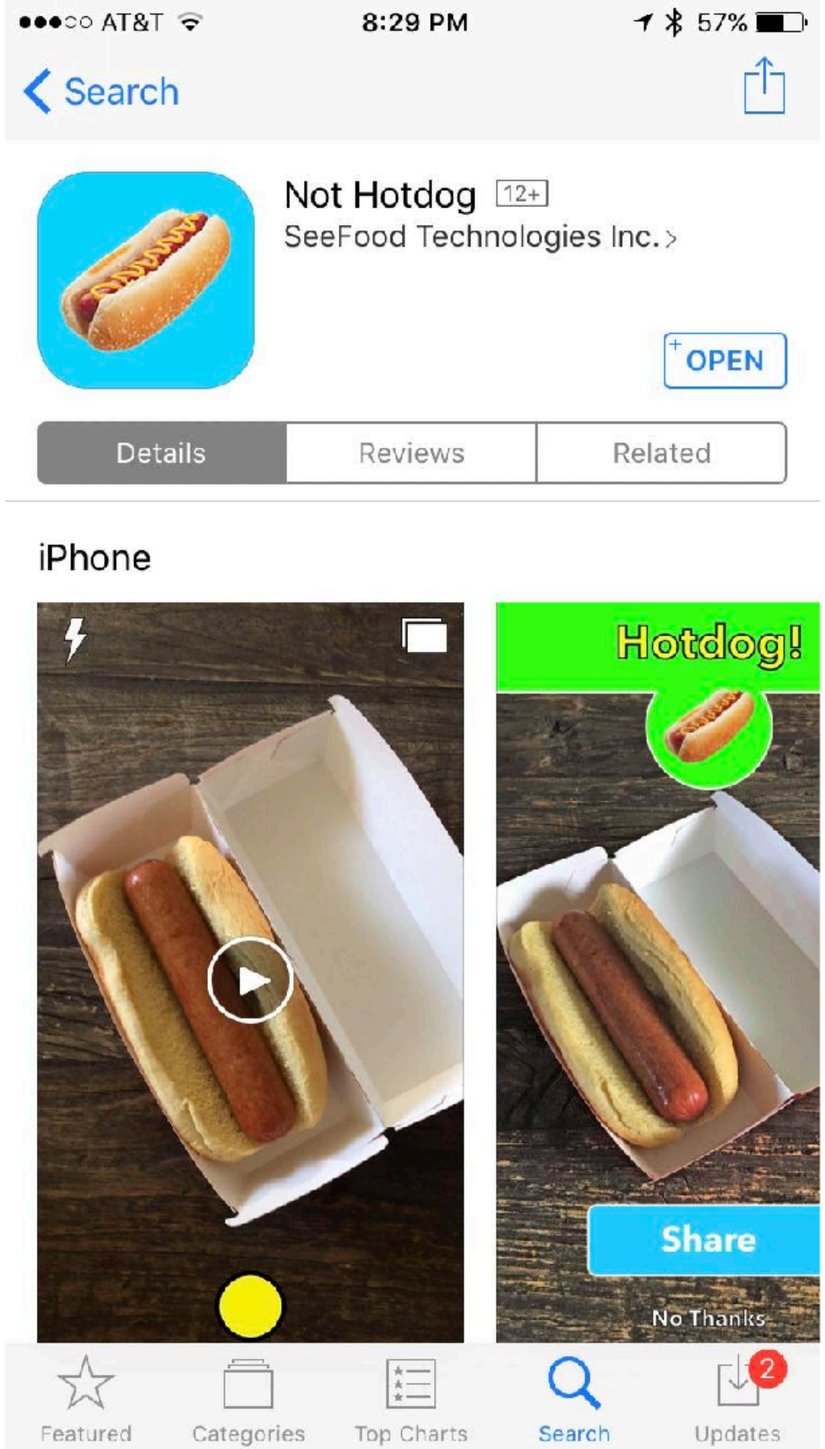


https://quickdraw.withgoogle.com/



https://www.youtube.com/watch?v=59bMh59JQDo

TRIFORK.
...think software



TRIFORK.
...think software

AT&T 8:29 PM 57%

Search

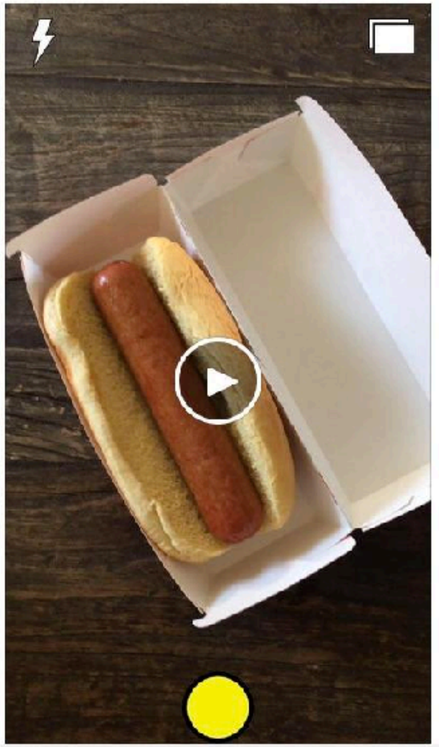
Not Hotdog 12+

SeeFood Technologies Inc.

OPEN

Details Reviews Related

iPhone



Featured Categories Top Charts Search Updates 2

●●●○○ AT&T 8:29 PM 57%

Search

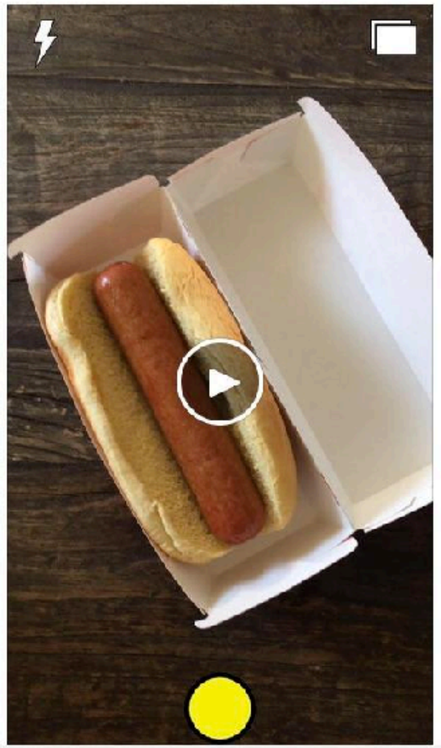
Not Hotdog 12+

SeeFood Technologies Inc. >

OPEN

Details | Reviews | Related

iPhone



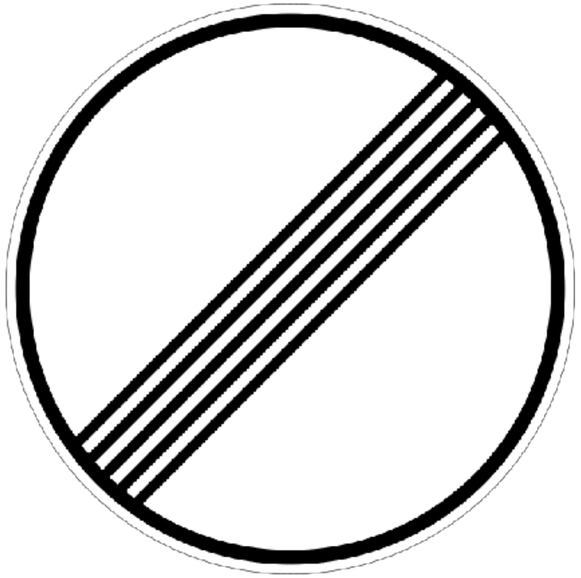
Featured Categories Top Charts Search Updates 2

http://www.tri.global/research/

TRIFORK.
...think software

https://en.wikipedia.org/wiki/Cycling_in_Copenhagen

TRIFORK.
...think software



https://en.wikipedia.org/wiki/Autobahn

TRIFORK.
...think software

Key-takeaways

TRIFORK.
...think software

Key-takeaways

1. Use MLaaS to get quick wins. Switch to Open Source when necessary

2. If a human can classify an image, then an ML model can as well

3. Be careful of introducing your own biases in your data

Thank You

Stefan Veis Pennerup
svp@trifork.com

TRIFORK.
...think software

goto; chicago

Please

Remember to rate this session

Thank you!

follow us @gotochgo

goto;
chicago

Click ‘Rate Session’

Rate **5** sessions to get the supercool GOTO reward

follow us @gotochgo