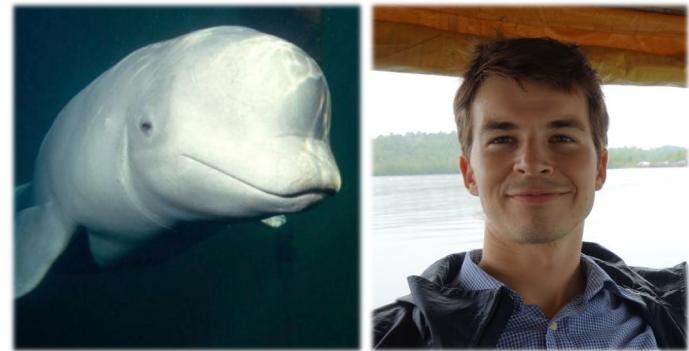


Kaggle Kernels

Why I think Kaggle is more than “just”
being the best data science competition platform

Gábor Fodor (beluga)



Meta Kaggle

Kaggle's public data on competitions, users, submission scores, and kernels



Kaggle Maintained by Kaggle • updated 10 hours ago

P Featured Prediction Competition

Quick, Draw! Doodle Recognition Challenge

How accurately can you identify a doodle?



\$25,000

Prize Money

G Google AI • 1,316 teams • 2 months ago

A Analytics Prediction Competition

NFL Punt Analytics Competition

Analyze NFL game data and suggest rules to improve player safety during punt plays

NFL The National Football League • 13 days ago



\$80,000

Prize Money



LVMH

Christian Dior
COUTURE

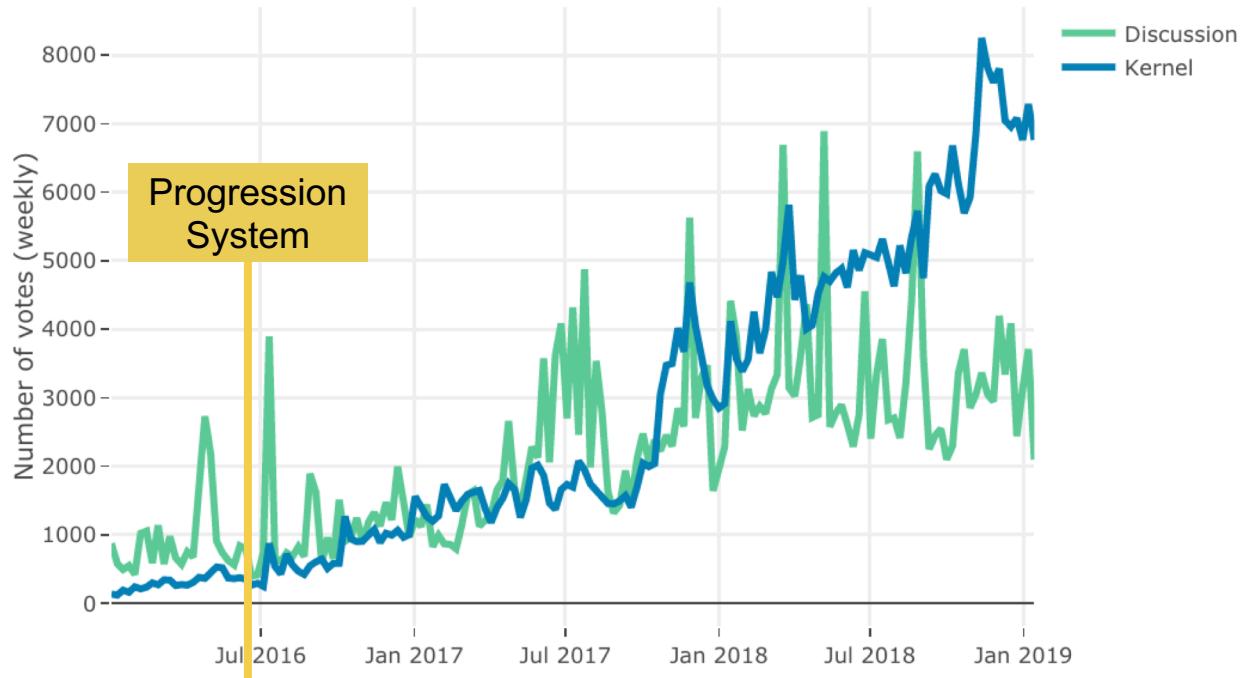
LOUIS VUITTON

SEPHORA

LOGICO

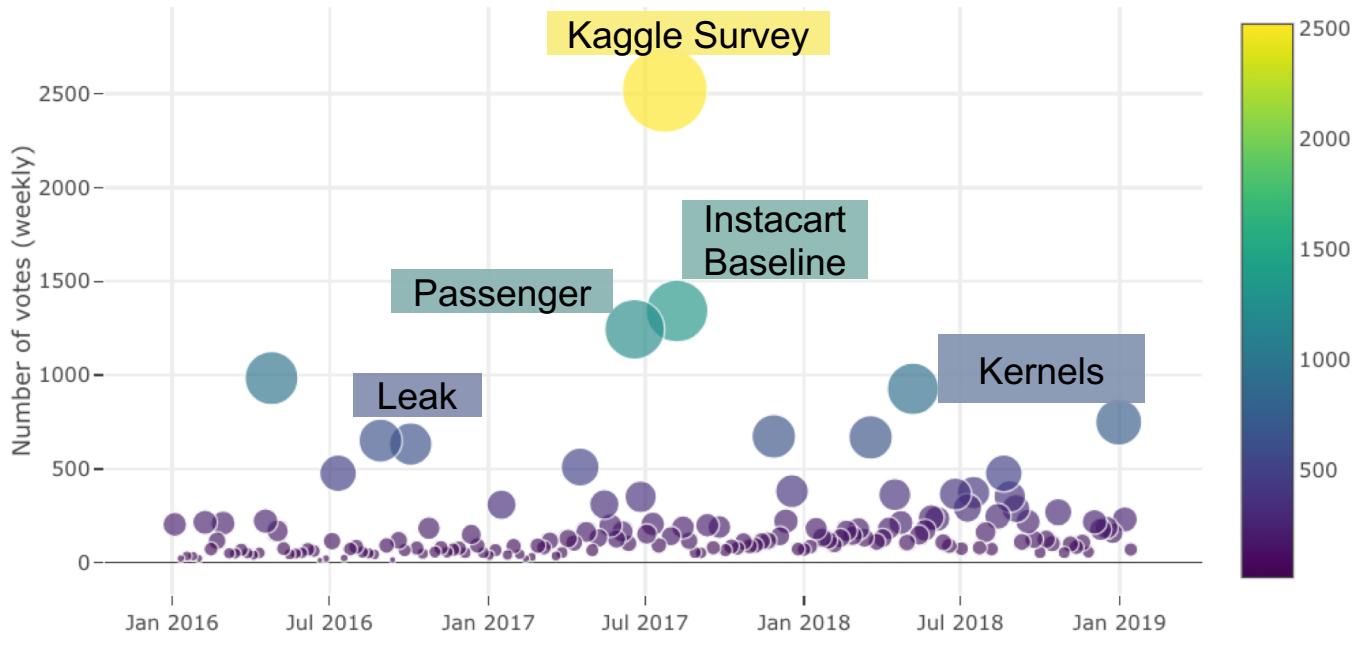
kaggle

Kernels are getting more popular





Hot topics



LVMH

Christian Dior
COUTURE

LOUIS VUITTON

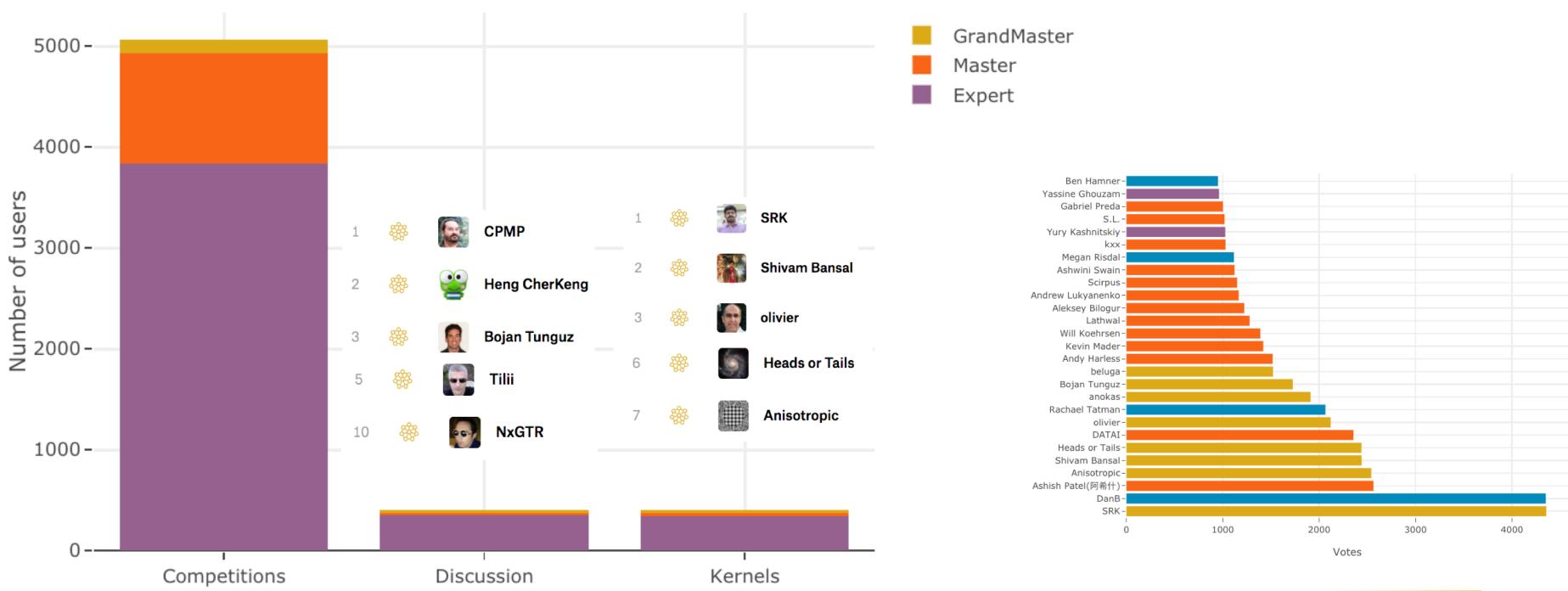
SEPHORA

LOGICO

kaggle



Discussion/Kernel Gap



LVMH

Christian Dior
COUTURE

LOUIS VUITTON

SEPHORA

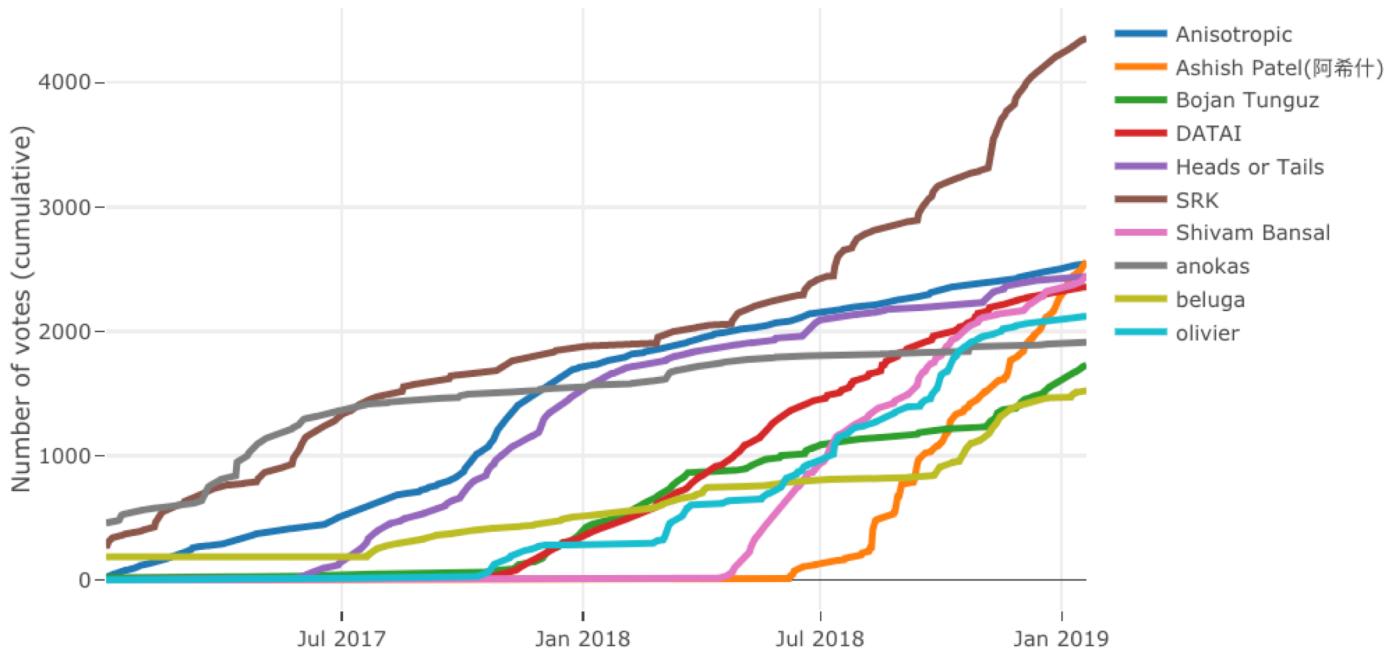
LOGICO

kaggle



Race for the top

kaggle
days



LVMH

Christian Dior
COUTURE

LOUIS VUITTON

SEPHORA

LOGICO

kaggle



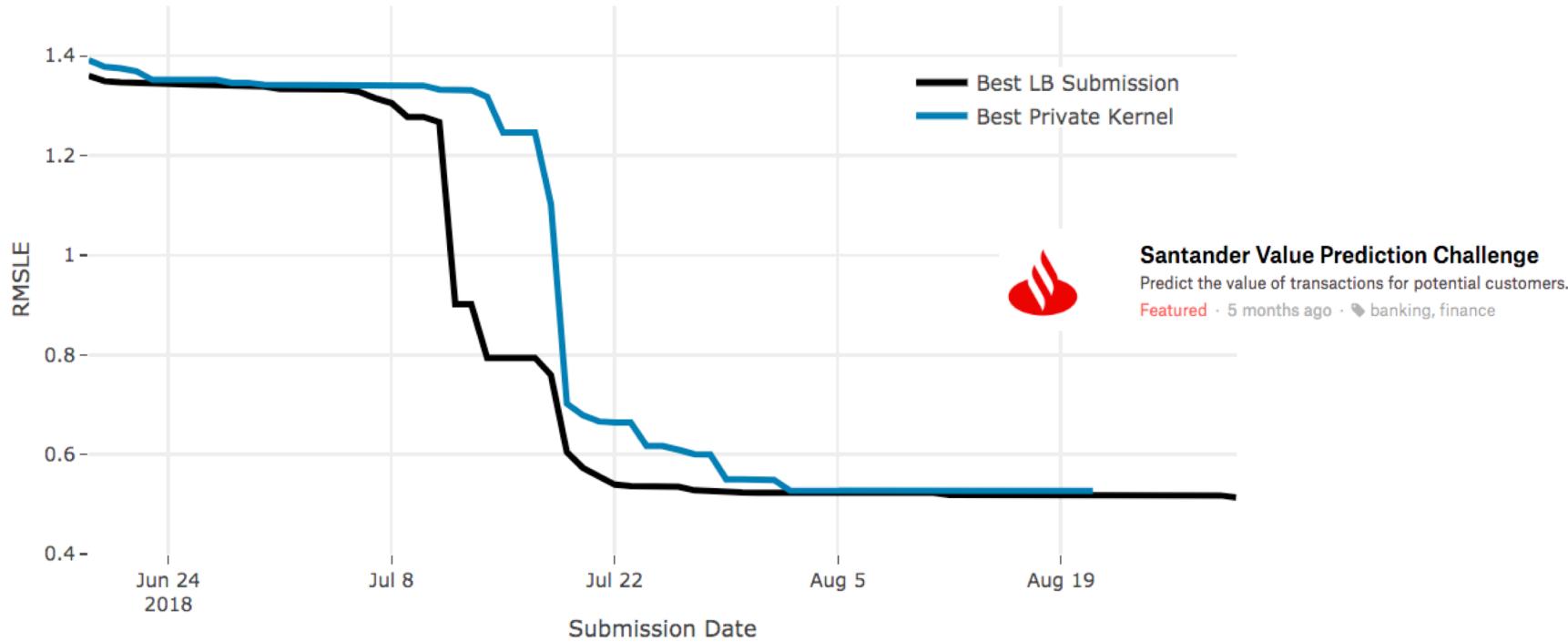
Random selected kernels



- [Pedro Marcelino: Comprehensive data exploration with Python](#)
- [William Cook & Keld Helsgaun: PMtest1](#)
- [Konstantin Lopuhin Mercari Golf: 0.3875 CV in 75 LOC, 1900 s](#)
- [Heads or Tails: Be my guest - Recruit Restaurant EDA](#)
- [Gilberto Titeritz: The Property by Giba](#)



Risk of oversharing



LVMH

Christian Dior
COUTURE

LOUIS VUITTON

SEPHORA

LOGICO

kaggle

LEVELS OF HELL:

LIKES KERNEL BUT DOESN'T UPVOTE

PRIVILEGELY FORKS KERNEL AND DOESN'T UPVOTE ORIGINAL

Gluttony

MAKES NEW KERNEL THAT JUST AVERAGES OTHER KERNELS WITH NO CREDIT

MAKES "BLEND OF BLENDS"

Heresy

MAKES "BLEND OF BLEND OF BLENDS"

MAKES "BLEND OF BLEND OF BLEND OF BLENDS"

FORKS, TWEAKS ONE HYPERPARAM, CLAIMS IT AS OWN WORK

FORKS, IMPROVES NOTHING, REMOVES CREDIT, BEGS FOR UPVOTES

@peterhurford

imgflip.com



LVMH

Christian Dior
COUTURE

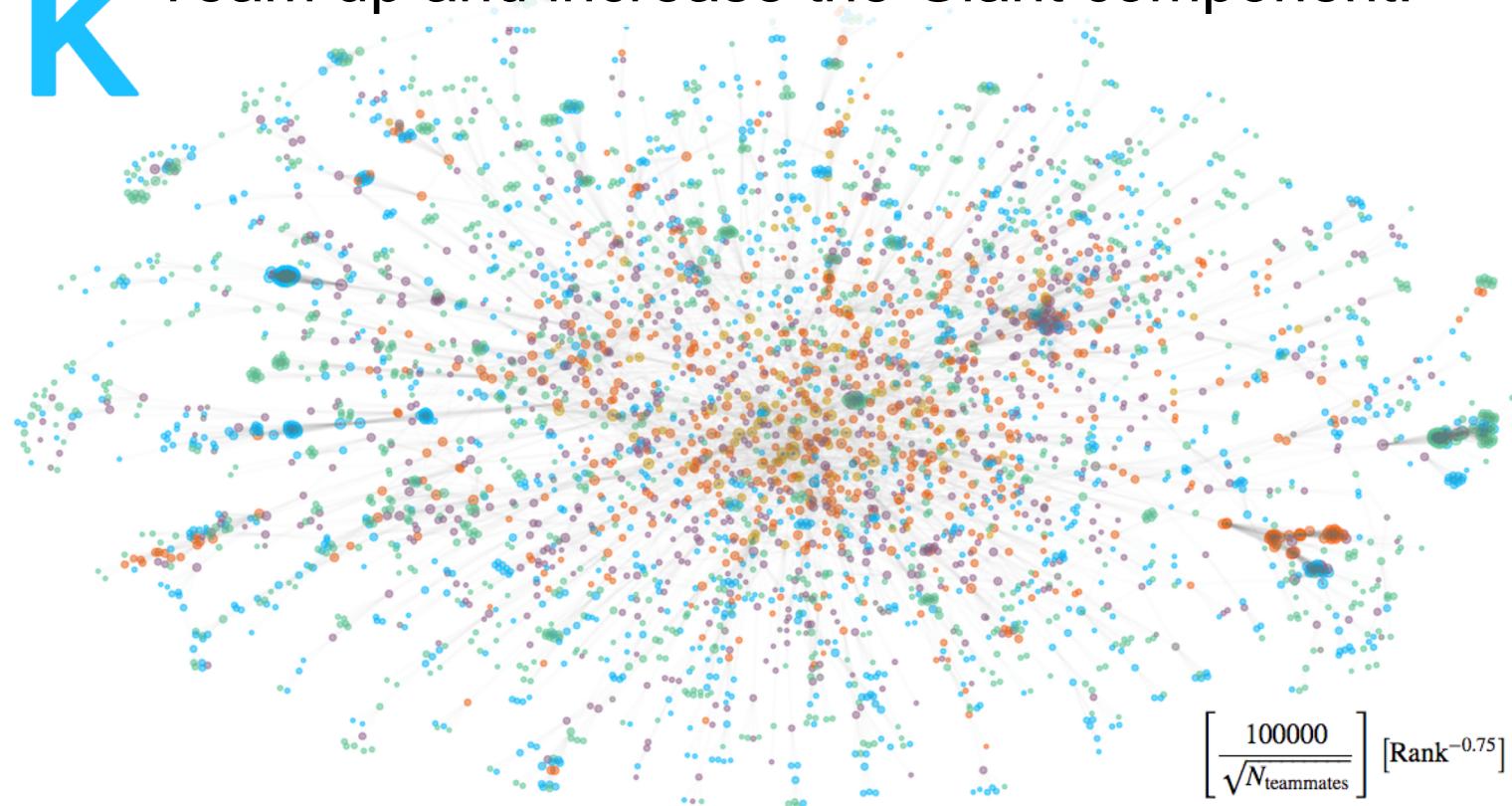
LOUIS VUITTON

SEPHORA

LOGICO

kaggle

k Team up and increase the Giant component!



$$\left[\frac{100000}{\sqrt{N_{\text{teammates}}}} \right] [\text{Rank}^{-0.75}] [\log_{10}(1 + \log_{10}(N_{\text{teams}}))] [e^{-t/500}]$$

LVMH

Christian Dior
COUTURE

LOUIS VUITTON

SEPHORA

LOGICO

kaggle



How to draw an owl?



50M drawings
340 categories

1st Pavel & Pavel ~ 0.95 MAP@3



LVMH

Christian Dior
COUTURE

LOUIS VUITTON

SEPHORA

LOGICO

kaggle



Computation Limits



9 hours execution time

5 GB output

16 GB temp

100 GB input

Network Access

Chain of kernels

External datasets

CPU

4 CPU cores

17 Gigabytes of RAM

GPU

NVIDIA Tesla K80

2 CPU cores

14 Gigabytes of RAM



252



- Greyscale MobileNet [LB=0.892]
2mo ago ⚡ 0.892 🔍 classification, deep learning, starter code

132



- Shuffle CSVs
2mo ago

89



- How to draw an owl? [LB=0.002]
2mo ago ⚡ 0.002 🔍 drawing, tutorial, data visualization, deep learning, starter code

40



- Black&White CNN [LB=0.77]
3mo ago ⚡ 0.771 🔍 tutorial, classification, deep learning, starter code

28



- (Un)recognized drawings ✅ ❌
3mo ago 🔍 beginner, data visualization, starter code

LVMH

Christian Dior
COUTURE

LOUIS VUITTON

SEPHORA

LOGICO

kaggle



🐘 Greyscale MobileNet



E2E

- Sample and shuffle
- Cv2 64x64 images
- Greyscale temporal encoding
- Image generator
- Keras MobileNet
- Local validation MAP@3
- Create Submission



LVMH

Christian Dior
COUTURE

LOUIS VUITTON

SEPHORA

LOGICO

kaggle



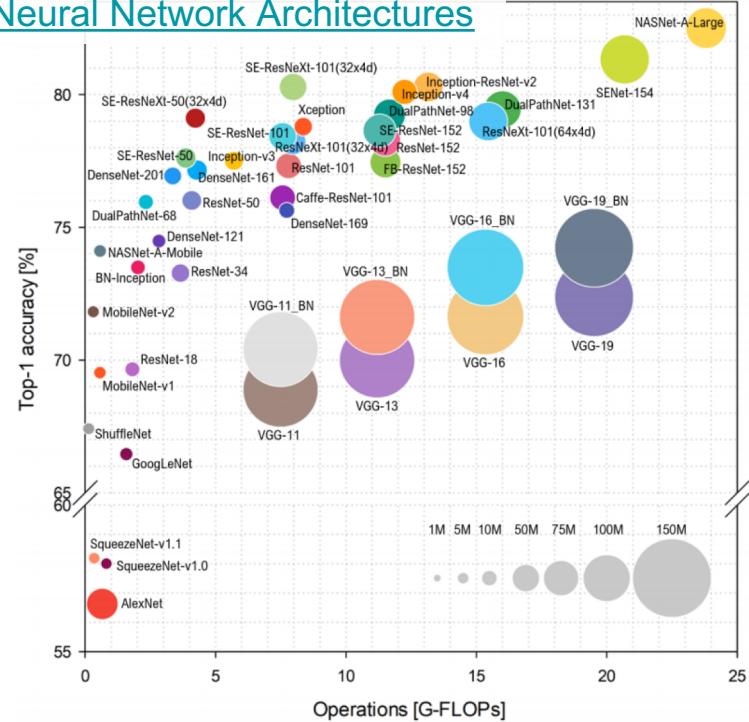
Architecture?



Benchmark Analysis of Representative Deep Neural Network Architectures

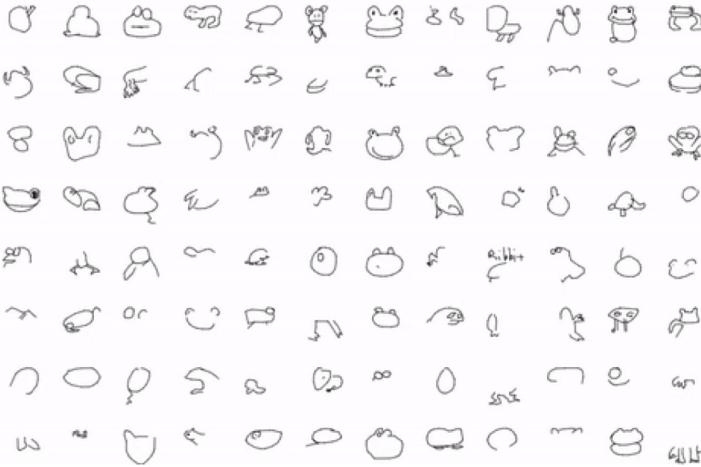
1st place solution

resnet18, resnet34, resnet50, resnet101, resnet152, resnext50, resnext101, densenet121, densenet201, vgg11, pnasnet, incresnet, polynet, nasnetmobile, senet154, seresnet50, seresnext50, seresnext101





CNN vs LSTM



[QuickDraw Baseline LSTM Reading and Submission](#)
[QuickDraw with WaveNet Classifier](#)
[Bidirectional LSTM Using Data Generator LB 0.825](#)

LVMH

Christian Dior
COUTURE

LOUIS VUITTON

SEPHORA

LOGICO

kaggle



Open ended Kaggle problems



Dataset Released Under CC0: Public Domain

786

Data Science for Good: Kiva Crowdfunding

Use Kernels to assess welfare of Kiva borrowers for \$30k in prizes

Dataset Released Under CC BY-SA 4.0

779

2018 Kaggle ML & DS Survey Challenge

Explore the 2018 Kaggle ML & Data Science Survey for \$28,000 in cash prizes

Analytics Prediction Competition

NFL Punt Analytics Competition

Analyze NFL game data and suggest rules to improve player safety during punt plays

\$80,000 Prize Money

The National Football League · 13 days ago

LVMH Christian Dior LOUIS VUITTON SEPHORA LOGICOOL kaggle

This image is a collage of three separate Kaggle competition banners. The top banner is for 'Data Science for Good: Kiva Crowdfunding', showing a landscape background and a count of 786 entries. The middle banner is for the '2018 Kaggle ML & DS Survey Challenge', featuring a woman working on a computer and a count of 779 entries. The bottom banner is for the 'NFL Punt Analytics Competition', showing a close-up of a football and a count of \$80,000 in prize money. All banners include the Kaggle logo and a timestamp of '13 days ago'. At the bottom of the collage, there are logos for LVMH, Christian Dior Couture, Louis Vuitton, Sephora, Logicool, and the word 'kaggle'.



Not a traditional kaggle competition



Propose specific rule modifications for the NFL
to reduce concussions during punt plays

Summary Slides & Kernels

Objective

- Solution Efficacy
- Game Integrity

Prizes

- 4x \$20K
- Pitch-Off > 2 Superbowl



Popularity vs Quality



37



• NFL Rules Wordcloud

 1mo ago  american football, sports, data visualization,

28



⊕ Summary - Budapest Pythons

 8h ago american football, sports, eda, data visualization



LVMH

Christian Dior
COUTURE

LOUIS VUITTON

SEPHORA

LOGICO

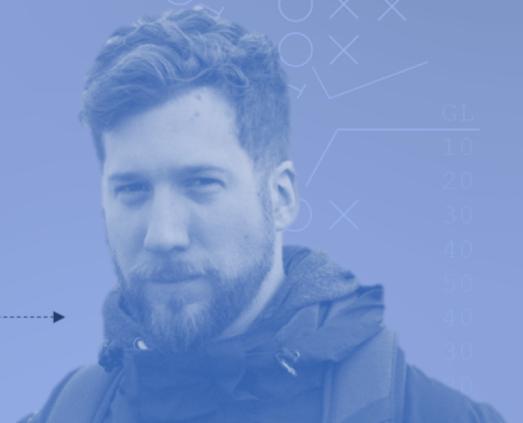
kaggle



BUDAPEST PYTHONS



ABOUT US
GABOR FODOR
L. KORTVELYESI



Summary Kernel

LVMH

Christian Dior
COUTURE

LOUIS VUITTON

SEPHORA

LOGICO

kaggle