

みんなのPython勉強会 #18

ODSC West参加報告

～シリコンバレーに行ってきました～

辻 真吾 (www.tsjshg.info)

2016.11.09

ちよつと宣伝

The screenshot shows the CTC Education Service website. At the top, there is a navigation bar with the CTC logo and the text "教育サービス" (Education Service). Below this is a search bar with a dropdown menu for "研修コース検索" (Search for training courses) and a search button. The main content area features a large banner for a column article titled "AI活用時代にPythonで見る夢" (Dreams seen in Python in the AI era) by Tsuji Tsutomu. The article is dated October 2016. To the right of the article, there is a sidebar with a section titled "関連するコース" (Related courses) listing Python-related courses. At the bottom right, there is a promotional banner for Cisco CCNA Routing and Switching, mentioning a limited-time campaign with Amazon.co.jp gift certificates.

CTC 教育サービス

これから学ぶ人も、資格取得を目指す人も、最適なカリキュラムを選べます。

カートを見る

研修コース検索

会場案内 受講までの流れ よくあるご質問 お問い合わせ

分野から探す ベンダーから探す プロダクトから探す 資格から探す レベルから探す

コース検索 例：分野名、ベンダー名、コースコードなど 検索

CTC教育サービス・ホーム > コラム > AI活用時代にPythonで見る夢 > 第1回 Pythonのすゝめ

コラム

AI活用時代に Python で見る夢

辻 真吾

CTC 教育サービス

[IT研修]注目キーワード [OpenStack](#) [OpenFlow/SDN](#) [情報セキュリティ](#) [Python](#) [システムトラブルシューティング](#)

第1回 Pythonのすゝめ (辻真吾)

2016年10月

こんにちは、辻真吾と申します。

ところでみなさんは、地球以外の惑星に生命が存在しているかどうか、考えたことがあるでしょうか？

私は、地球のような惑星が、この宇宙にはたくさんあると思っています。

ある惑星は、地球の恐竜時代かもしれませんし、別の惑星は、人類の祖先のような生き物が、洞穴で暮らし、よく磨いた石製の武器や 猿狩りに出掛けるとアスカもいれません。

関連するコース

- Pythonの関連コース一覧
- Python 入門1
- Python 入門2

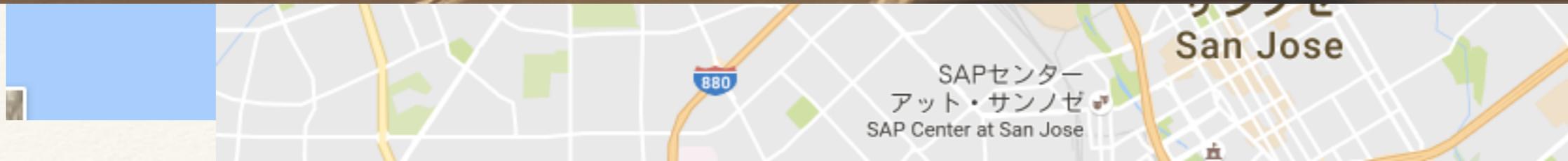
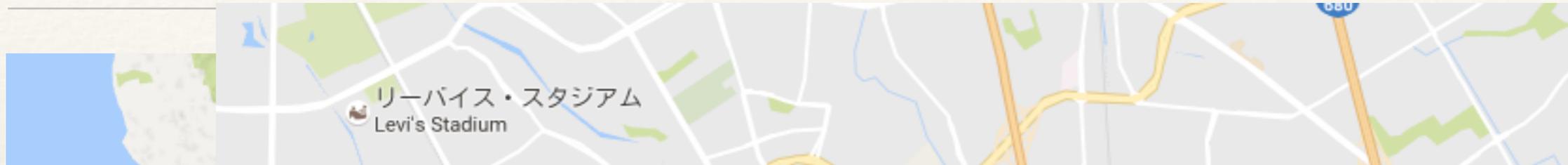
Cisco CCNA Routing and Switching 受講キャンペーン実施中!

2016/12 まで 期間限定キャンペーン!

受講者全員にもれなく amazon.co.jp キット券

Citrix

Santa Clara





Open Data Science Conference

- ❖ 創設者 Sheamus McGovern
- ❖ making data science more accessible
- ❖ オープンソースを使ったデータサイエンスを広げるための活動
- ❖ 北米を中心に世界各国で開催
- ❖ ODSC East (ボストン) が最大で4,000人以上の参加者
- ❖ 今回のODSC Westは2,000人強



<https://www.odsc.com/>

基調講演 (Keynotes)

はじまるよー。講演聞くのもいいけど友達作ってね。

PythonでもRでもいいけどオープンソースで行こう。でもAIが進化して、世界もそのうち終わるけどねー



Sheamus

McGovern

ODSC Founder and Chair



Bob Rogers

Chief Data Scientist: Analytics and Artificial Intelligence Solutions Group at Intel Corp.



Jeremy Achin

CEO and founder of DataRobot



Gary Smerdon

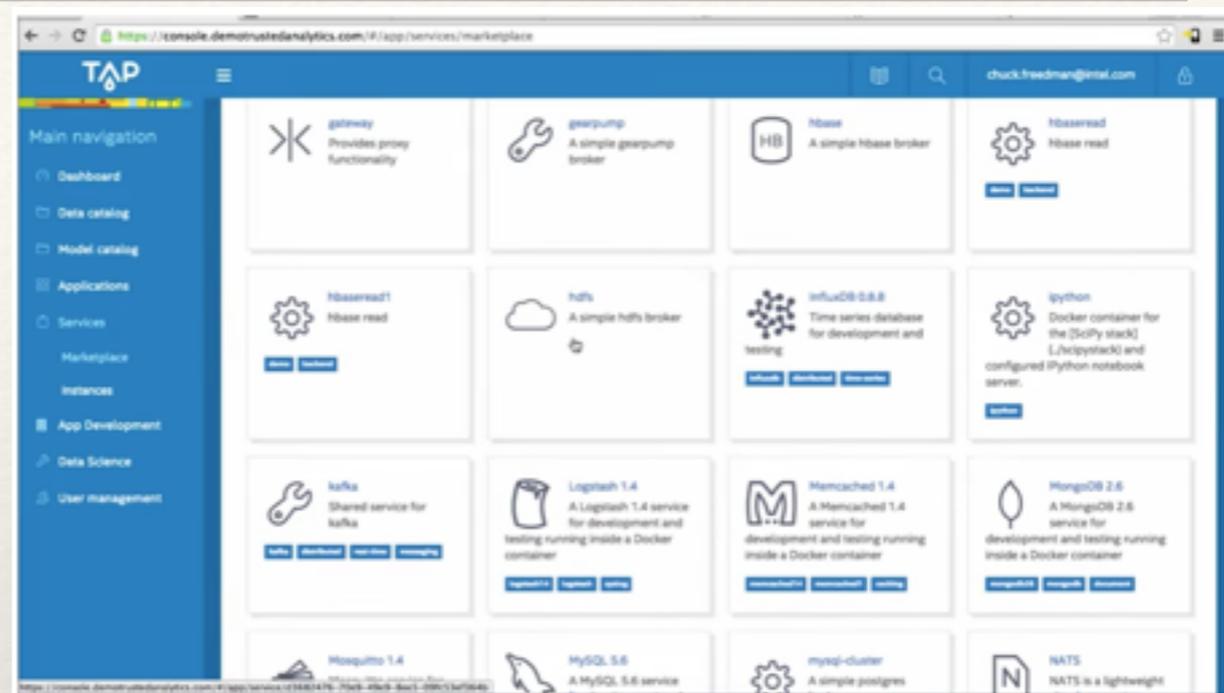
President & CEO at TidalScale

AIがデータサイエンスを変える時代。いろいろなデータを統合して解析できる環境が重要

Software-defined Serverが世界変えます。俺がGame changerだ！

Trusted Analytics Platform (TAP)

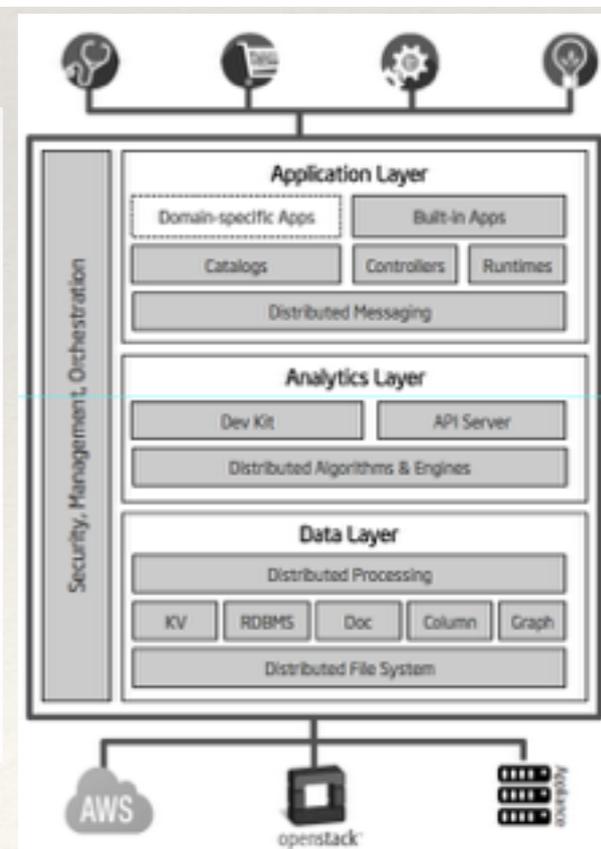
- ❖ データサイエンスのための統合的なプラットフォーム
- ❖ オープンソース
- ❖ データの取得から解析までのパイプラインを管理できる (?)
- ❖ ipynb形式のファイルを組み込むことも可能
- ❖ <http://trustedanalytics.org/>

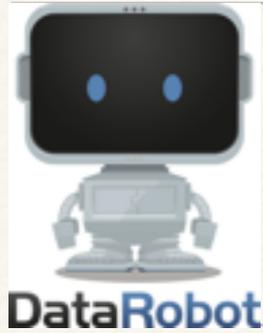


Data Science on TAP
Extensible tools, scalable algorithms and powerful engines to train and deploy predictive models

App Development on TAP
Polyglot platform for cloud-native applications with consistent APIs and easy to use services

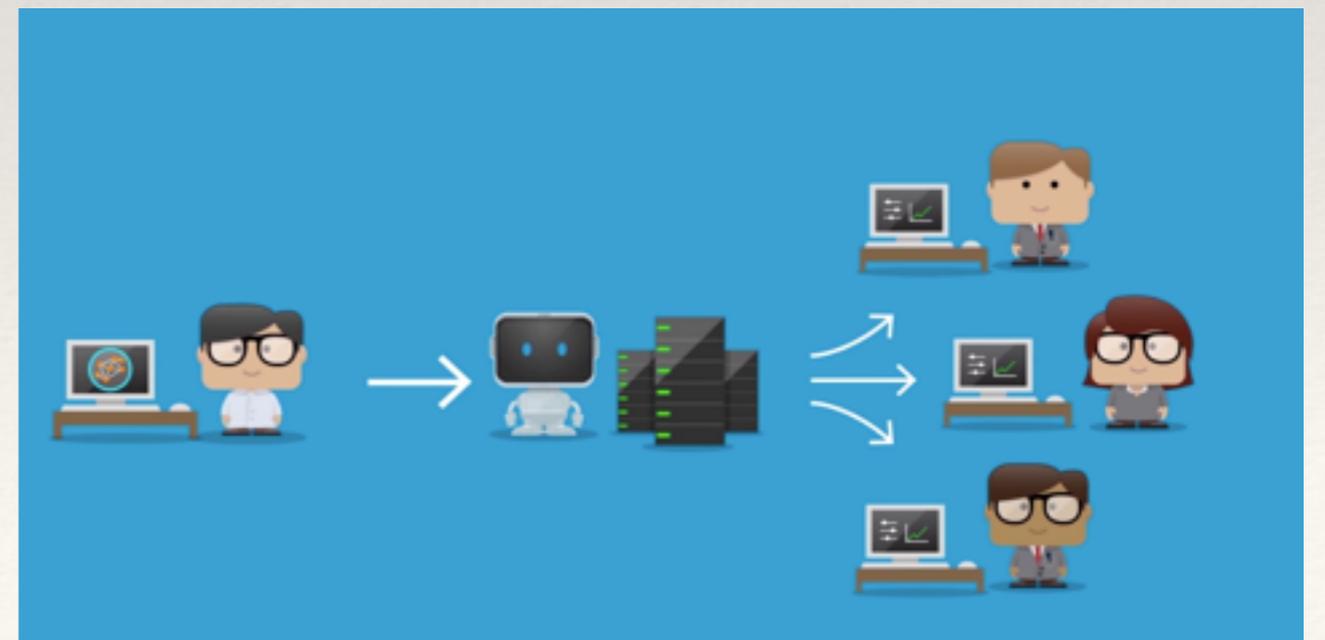
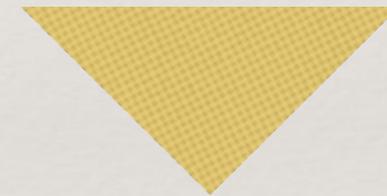
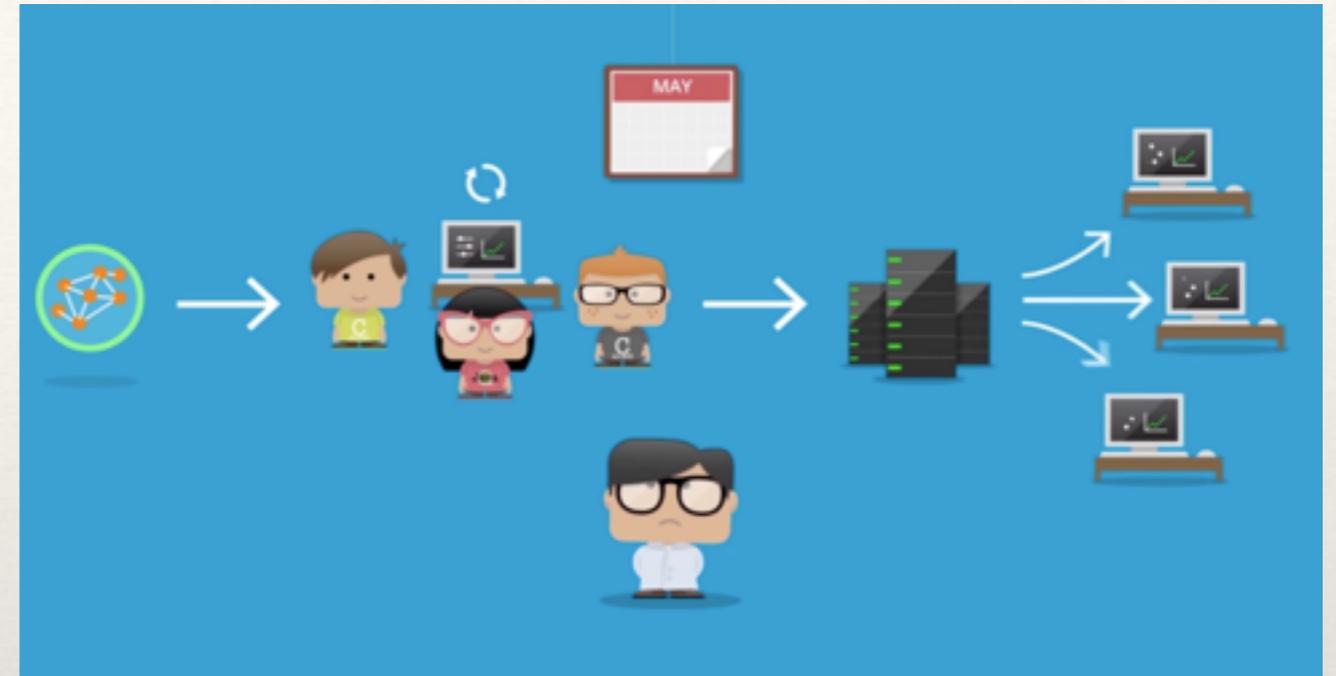
System Operations on TAP
Cross-cutting view of the entire platform with easy management and provisioning for private and public cloud infrastructure





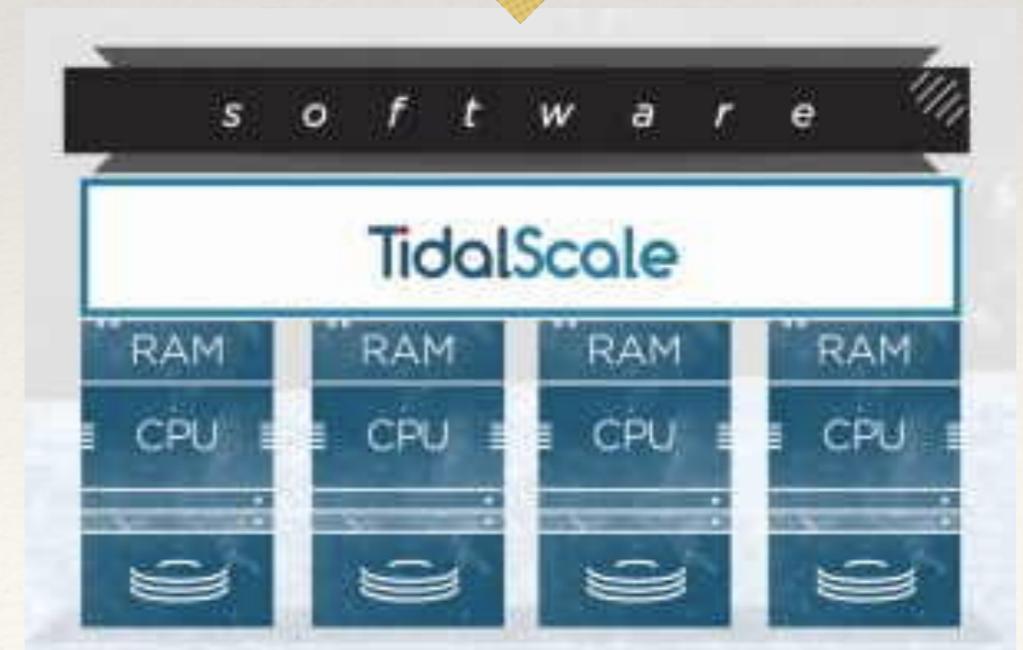
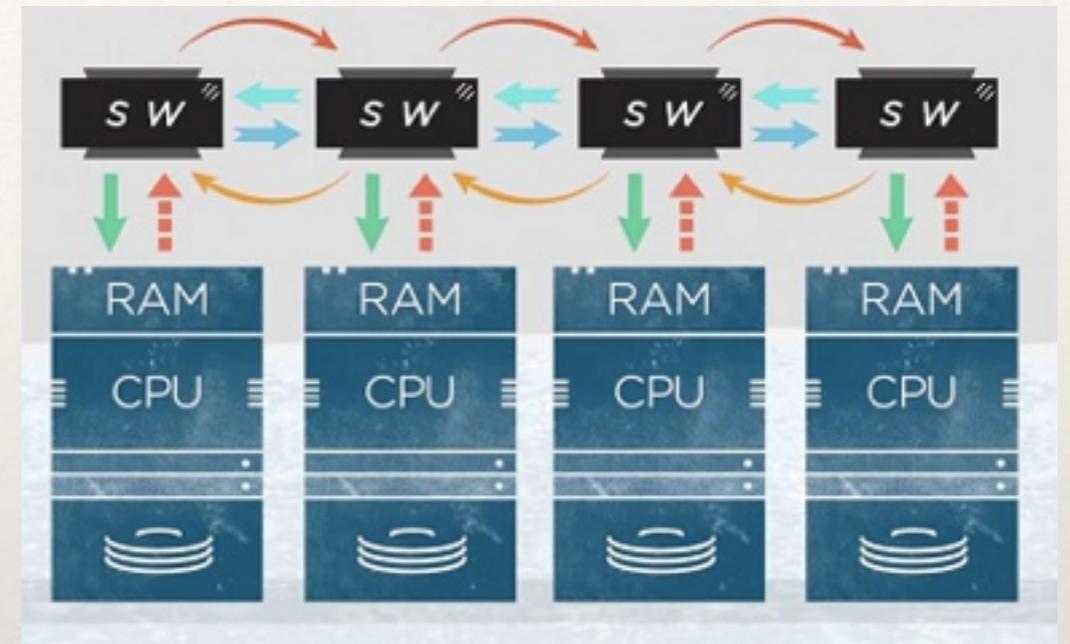
DataRobot

- ❖ 今回の最大スポンサー
- ❖ データサイエンティストの仕事
を自動化
- ❖ PythonやRで作ったモデルを
自動的に取捨選択？
- ❖ 日本は2番目の市場
 - ❖ リクルートテクノロジーズ
 - ❖ NS Solutions (販売代理)



TidalScale

- ❖ Software-defined server™
- ❖ データの量は増え続ける
- ❖ 年月が過ぎるとデータの価値は下がる
- ❖ 拡張性と柔軟性
 - ❖ seamless scalability
- ❖ TidalScaleの中のリソース配分には機械学習が使われている



Google

- ❖ Kazunori Sato
- ❖ 実際にDeep Learningを使っているのは、会場を埋めた300の約3割弱
- ❖ play ground
- ❖ TensorFlowの楽しい応用例
- ❖ TPU: TensorFlow Processing Unit

Machine Learning API

- Speech API
- Vision API
- Natural language API

Cloud Machine Learning

TensorFlow

社内でもこれがプロダクト開発に使われている

お手軽



専門家向け

TensorFlow playground

Iterations: 000,000
Learning rate: 0.03
Activation: Tanh
Regularization: None
Regularization rate: 0
Problem type: Classification

DATA: Which dataset do you want to use?
Ratio of training to test data: 50%
Noise: 0
Batch size: 10
REGENERATE

FEATURES: Which properties do you want to feed in?
 X_1
 X_2
 X_1^2
 X_2^2
 $X_1 X_2$
 $\sin(X_1)$
 $\sin(X_2)$

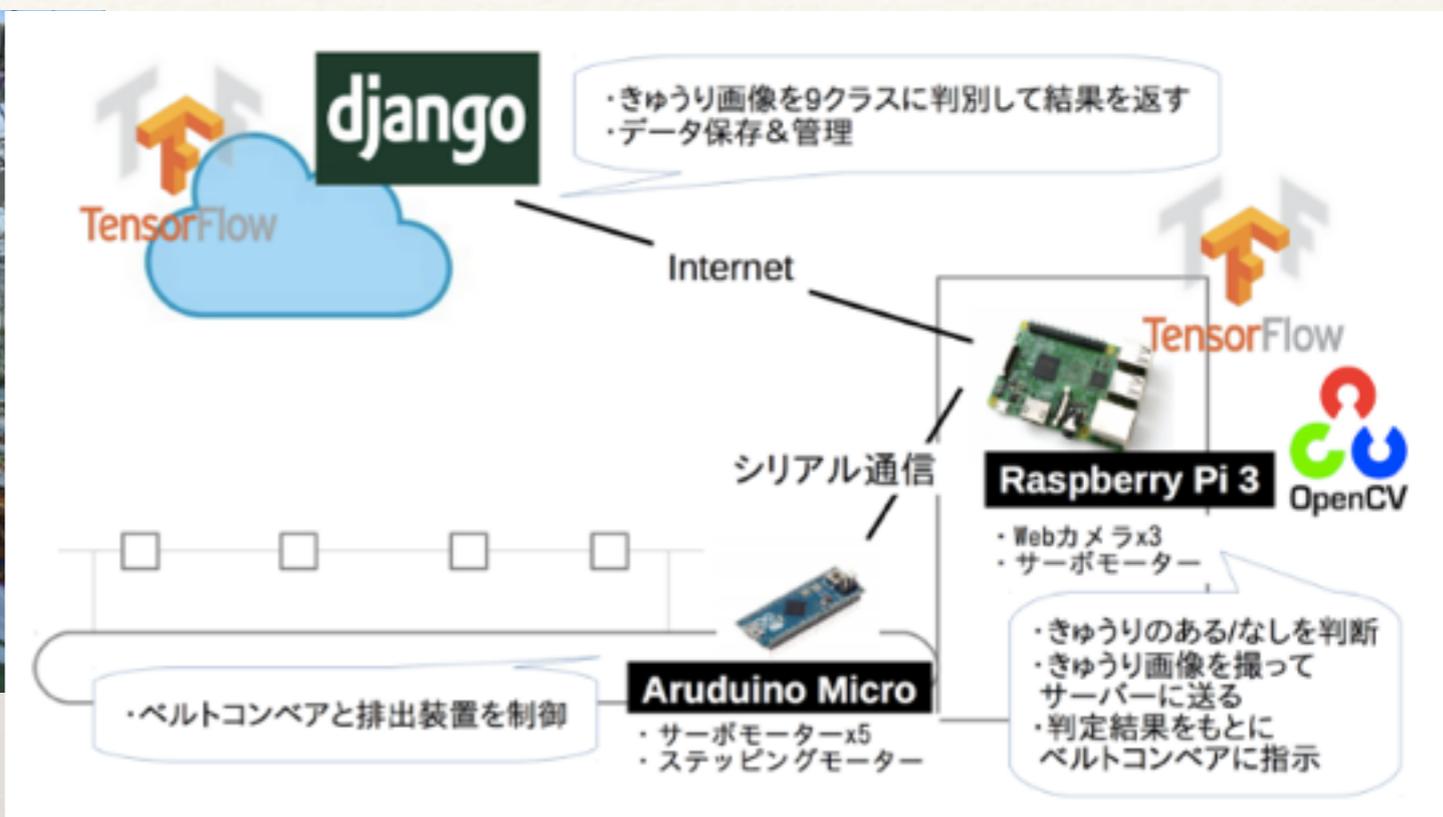
2 HIDDEN LAYERS
4 neurons
2 neurons

The outputs are mixed with varying **weights**, shown by the thickness of the lines.
This is the output from one **neuron**. Hover to see it larger.

OUTPUT: Test loss 0.506
Training loss 0.509

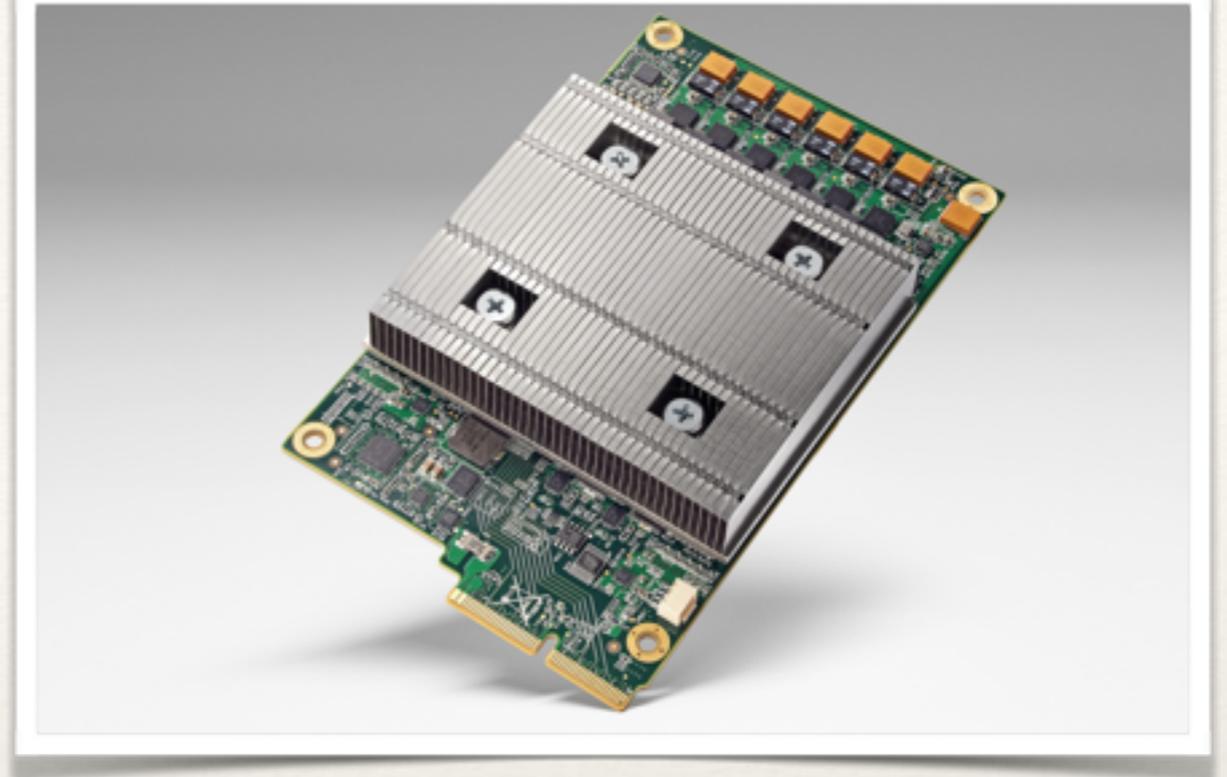
Colors shows data, neuron and weight values.
 Show test data Discretize output

キュウリの仕分け



元組み込みエンジニアの息子さんが、
キュウリの仕分けという人の仕事を
人工知能で奪った事例*

*お母さんは喜んでいましたそうです。



- ❖ CPUでもGPUでもなくTPU
 - ❖ TensorFlow Processing Unit
- ❖ TensorFlow専用のチップ
- ❖ ムーアの法則が終わった今、ハードウェア進化の1つの方向性

Kaggle

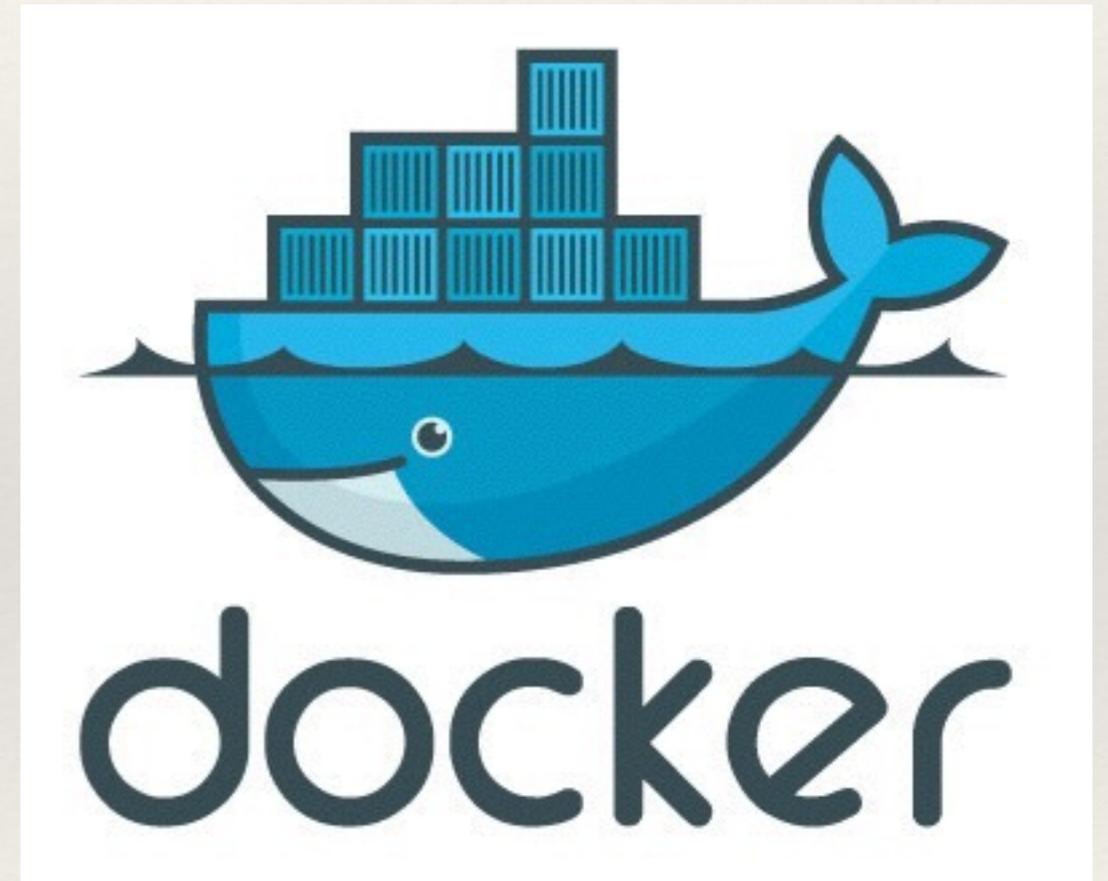
- ❖ 機械学習モデルのコンペティションサイト
- ❖ 比較的簡単な分類 (Classification) の問題
 - ❖ 募集24時間で、0.826173の精度ができるがその後数ヶ月で0.829072までしか伸びない
 - ❖ Scikit-learnなどAPIが充実しているので、誰がやっても同じ時代になりつつある

Kaggle 続き

- ❖ 難しいのは、例えばQ&A
 - ❖ 自然言語の質問に対して4択から正解を選ぶ
 - ❖ 43.3%だった当初の精度が、59.3%まで改善
 - ❖ 最後まで、全参加者が半々で選ぶ選択肢があるなど、まだまだ発展途上

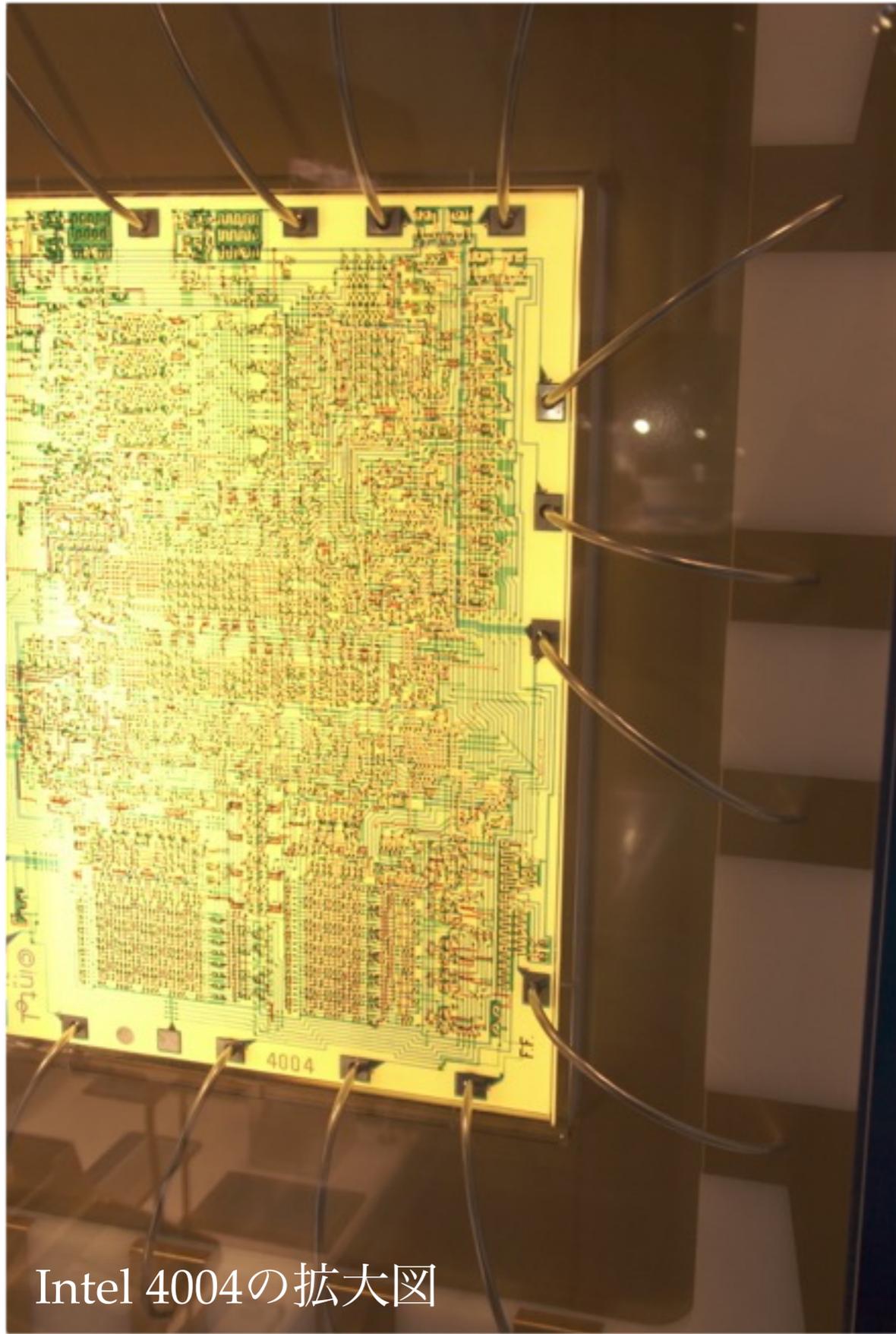
再現性のあるモデル作り

- ❖ 再現性のある機械学習もできるためには、すべてをバージョン管理する必要がある
 - ❖ code
 - ❖ data
 - ❖ environment
- ❖ Dockerのimageとして環境を提供する試み
 - ❖ Kaggle Kernels



おまけ

Intel museumに行ってみた



Intel 4004の拡大図



本社の入口



“Optimism is an essential ingredient of innovation. How else can the individual welcome change over security, adventure over staying in safe places?”

–Robert Noyce (1927-1990)

革新的なことをするには、楽観的に考えることが絶対に必要なんだ。
そうでなければ、誰も、危険を冒してまで旅に出たりはしないだろ？

まとめ

- ❖ 大手企業がオープンソースを提供
- ❖ ムーアの法則の終わり
- ❖ 機械学習のお手軽API化
- ❖ 再現性 (reproducibility) のある解析
 - ❖ TAP, Kaggle kernelsなどの試み
- ❖ 自然言語処理、形の違うデータの統合などが今後の課題

ありがとうございました。