

AIエンジニア研修 ディープラーニングの基礎



- 1 ニューラルネットワーク
- 2 ディープラーニングの概要
- 3 活用事例
- 4 ライブラリとフレームワーク

演習内容

Kerasによる手書き文字認識

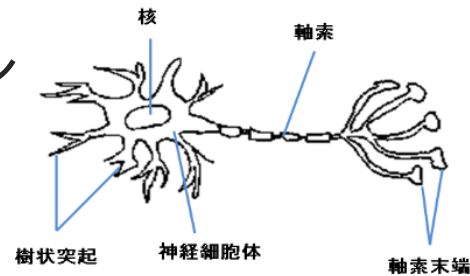
1. ニューラルネットワーク

ニューラルネットワークとは

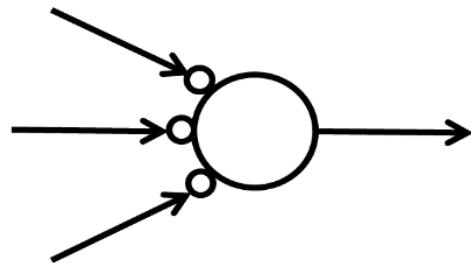
- 生物の脳のニューロンを模した機械学習の一種
ニューロンに見立てた変数を多数連結して作ったもの
- これまで3回ほどブームと言われた時期があったが
ディープラーニングの発展で2012年から再注目されている

入力 → 出力

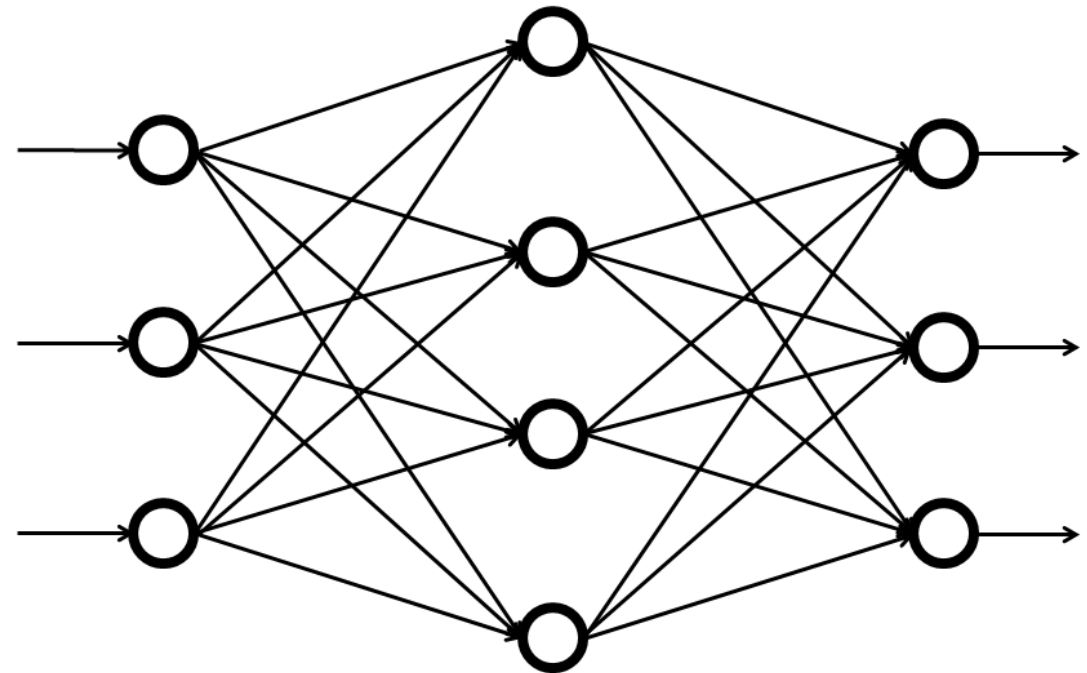
生物のニューロン



人工ニューロン



ニューラルネットワーク



ニューラルネットワークブーム

第一次NNWブーム

パーセプトロン



1950年代

1960年代

1970年代

1980年代

1990年代

2000年代

2010年代



第一次AIブーム

推論と探索



第二次AIブーム

エキスパートシステム



第二次NNWブーム

多層パーセプトロン



第三次NNWブーム

ディープラーニング



第三次AIブーム

機械学習

第一次ブーム | パーセプトロンの発見

- 生物の神経細胞の発火モデル（McCulloch-Pittsモデル）が提案された（1943年）
- シナプス可塑性に関する仮説（ヘップ則）提唱（1949年）
 - ニューロンAの発火がニューロンBを発火させると2つのニューロン間の結合が強まる
 - 現在のニューラルネットワーク学習の基本的な考え方
- ローズンブラットが単純パーセプトロンを考案（1958年）
 - 2クラス分類モデル
 - 非線形問題が解けないことをミンスキーらが指摘、ニューラルブーム自体が下火に

ウォルター・ピッツ
(1923-1969)



数学者。
ニューラルネットワークの元祖とも言える
「形式ニューロン」を
考案した。

▶カリフォルニア州立大学より

マービン・ミンスキー
(1927-2016)



認知科学者。
MITのAI研究所や
ダートマス会議の創
設者でAIという言葉
を生み出した。

Wikipediaより

ジョン・マッカーシー
(1927-2011)



計算機科学者。
ミンスキーと並び
AIの研究分野自体
の創出に尽力した
AI界の巨人。

Wikipediaより

第二次ブーム | 多層パーセプトロンへの期待と失望

- パーセプトロンを**多層化**すれば非線形問題(XOR問題)を解けることが示される (1986年)
 - AND、NAND、ORゲートを組み合わせることで実現
- 多層パーセプトロンを効率よく学習できる誤差逆伝播法の考案 (1986年)
 - **学習データの量、計算リソースの問題**から当時は3層程度が限界とされていた
 - この規模で学習できることは限られておりニューラルネットワーク研究は再び冬の時代へ

福島邦彦
(1936-)



計算機科学者。
1979年に畳み込み型の多層パーセプトロン「ネオコグニトロン」を考案し、現代のディープラーニングに影響を与える。

▶ 福島邦彦より

ジェフリー・ヒントン
(1947-)



計算機科学者。
ニューラルネットワーク研究で重要な技術である「誤差逆伝播法」「ボルツマンマシン」「自己符号化器」「DBN」などの発明に関わってきたディープラーニング研究の第一人者。
現在はGoogleにも所属している。

▶ Wikipediaより

第三次ブーム | 2000年代後半～現在

- 90年代から2000年代前半：低迷期だが研究は進んでいた
 - ユニットの振る舞いを確率的に記述した**ディープビリーフネット（DBN）**が注目を集める
 - 入力そのものを出力するように学習する**自己符号化器**でも事前学習できると明らかに
- 2000年代後半：音声認識や画像認識のベンチマークテストで多層ニューラルネットが記録を更新するように
 - 要因は主に次の3つと言われている
 - ・ ウェブから**大量の学習用データ**の収集が可能に
 - ・ GPUやマルチコアCPUによる**並列計算能力**のコストパフォーマンス向上
 - ・ 自己符号化器や畳込みニューラルネットワークなどの**研究の進展**
- 深層学習（ディープラーニング）のブームへ



Nvidia社製GPU Tesla
ハイエンド**CPUの10倍以上**
の計算性能を持つコア数は
100倍以上

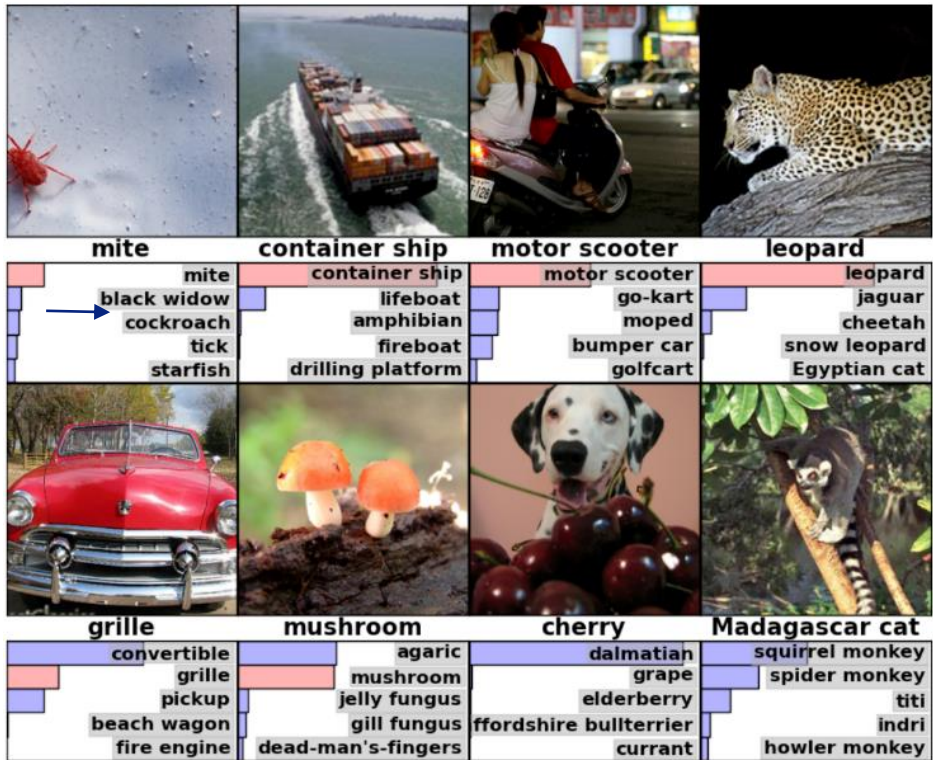
ディープラーニングが画像認識コンテストで優勝

画像分類の国際コンテストであるILSVRC-2012で
トロント大学ヒントン教授のグループがclassification
とlocalizationの部門で大差をつけて優勝

classification : 与えられた画像を1000カテゴリに分類する
localization : 分類に加え画像中の物体の位置を推定する

classificationの例

与えられた各画像について、そのカテゴリの候補を5つ予測し、その中に正しいカテゴリが含まれているかを評価



►Kriahevsky et al. 2012 (一部抜粋)

classification

Team name	Error (5 guesses)
SuperVision	0.15315
ISI	0.26172
OXFORD-VGG	0.26979
XRCE/INRIA	0.27058

localization

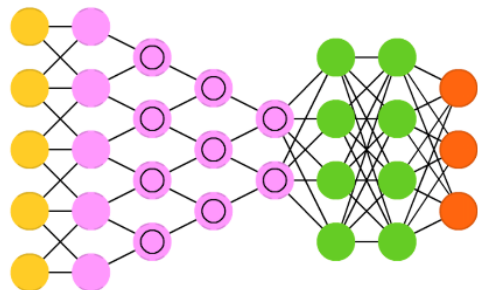
Team name	Error (5 guesses)
SuperVision	0.335463
OXFORD-VGG	0.500342
ISI	0.536474

ニューラルネットワークの種類

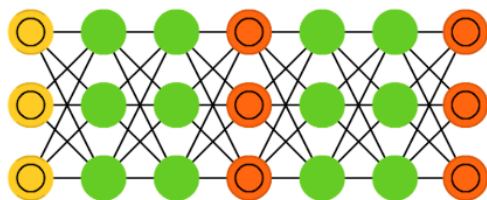
- ・ **教師あり学習**の機械学習（自己符号化器など一部は教師なし）
- ・ 左から右に一方通行で計算を行っていく **順伝播型**
- ・ 逆方向への経路を持つ **再帰型**

順伝播型（フィードフォワード）

畳み込みニューラルネット

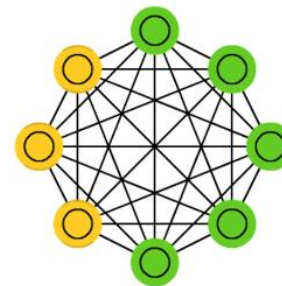


Generative Adversarial Networks



再帰型（リカレント）

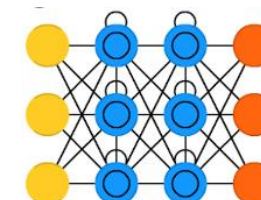
ボルツマンマシン



自己符号化器



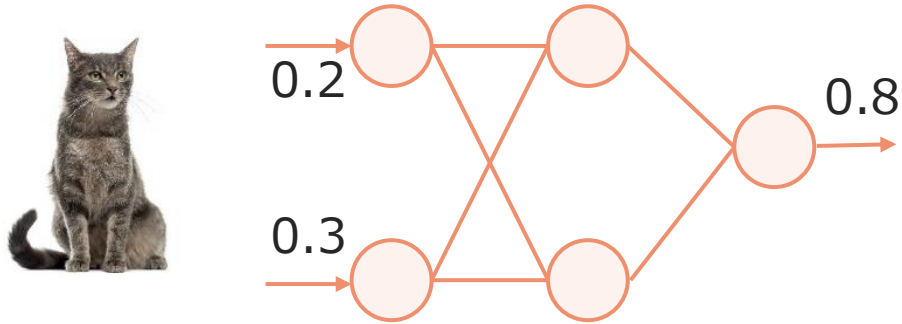
LSTMネットワーク



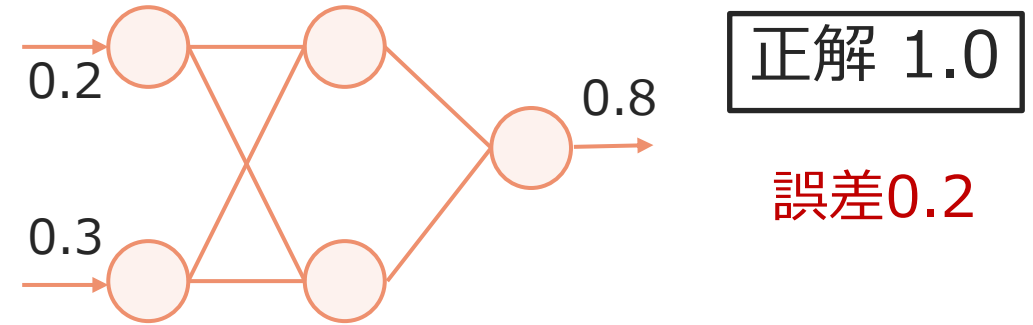
▶ 画像の出典 (2017.09.04) <http://www.asimovinstitute.org/neural-network-zoo/>

ニューラルネットワーク学習の流れ

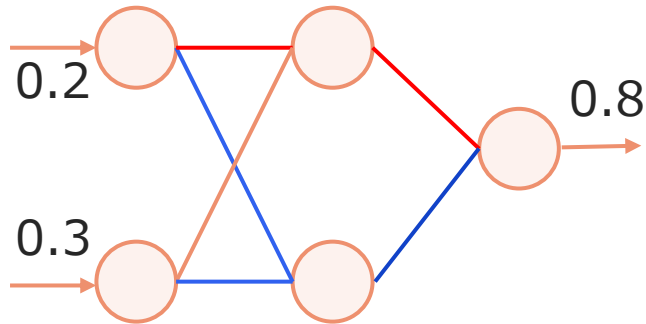
①入力データを元に出力を計算する



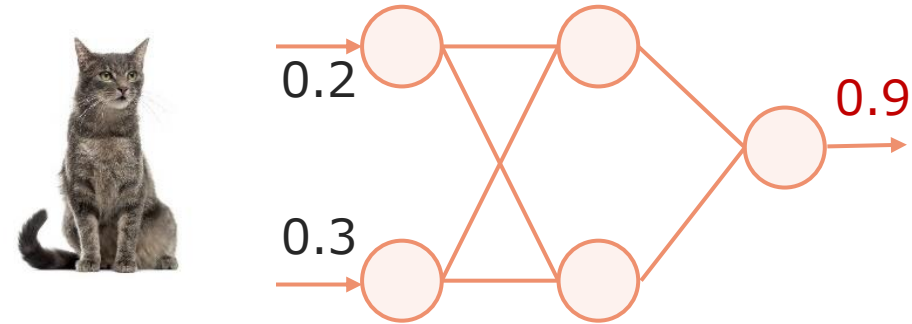
②正解との誤差を計算する



③誤差が小さくなるよう重みを順に更新



④重みが更新されなくなる（収束する）
まで①～③を繰り返す



参考 | 勾配降下法

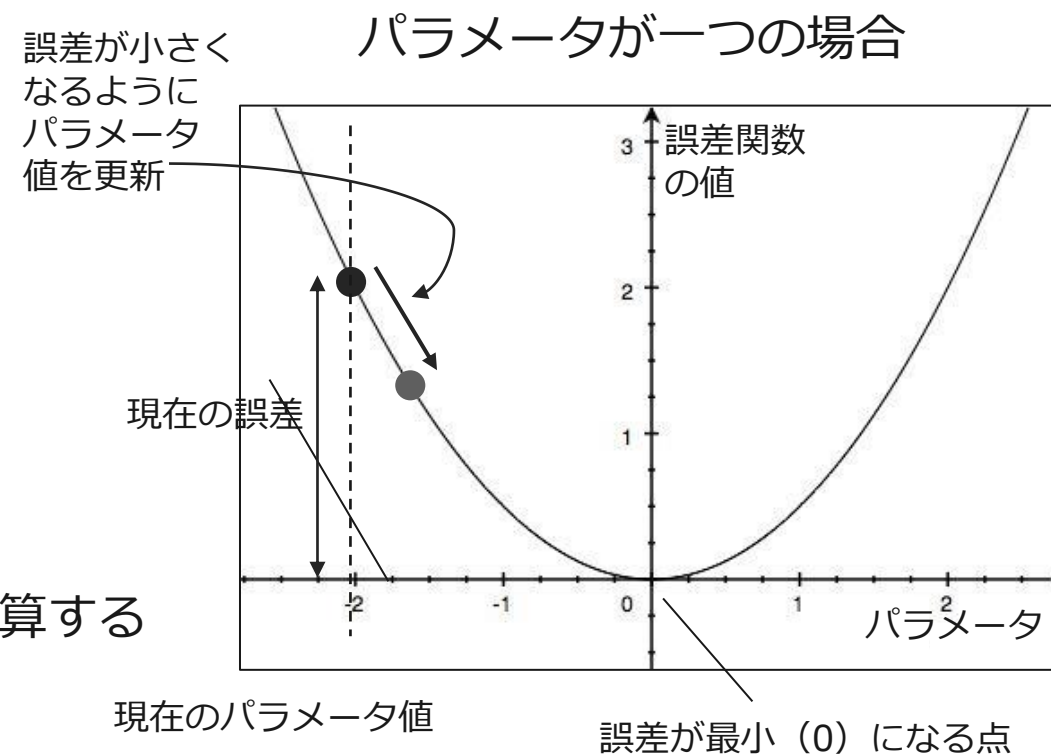
- 関数の勾配（傾き）を**降下する方へパラメータを繰り返し更新**することで関数の最小値（あるいはできるだ小さな値）を探索するアルゴリズム
- 誤差関数の値が小さくなるようにネットワークの重みとバイアスを少しずつ更新していくことで、ネットワークの出力を正解出力に近づける

重みの更新式

$$\underset{\text{更新後の重み行列}}{W^{(t+1)}} = \underset{\text{現在の重み行列}}{W^{(t)}} - \underset{\text{学習率}}{\eta} \underset{\text{誤差関数の勾配}}{\nabla E}$$

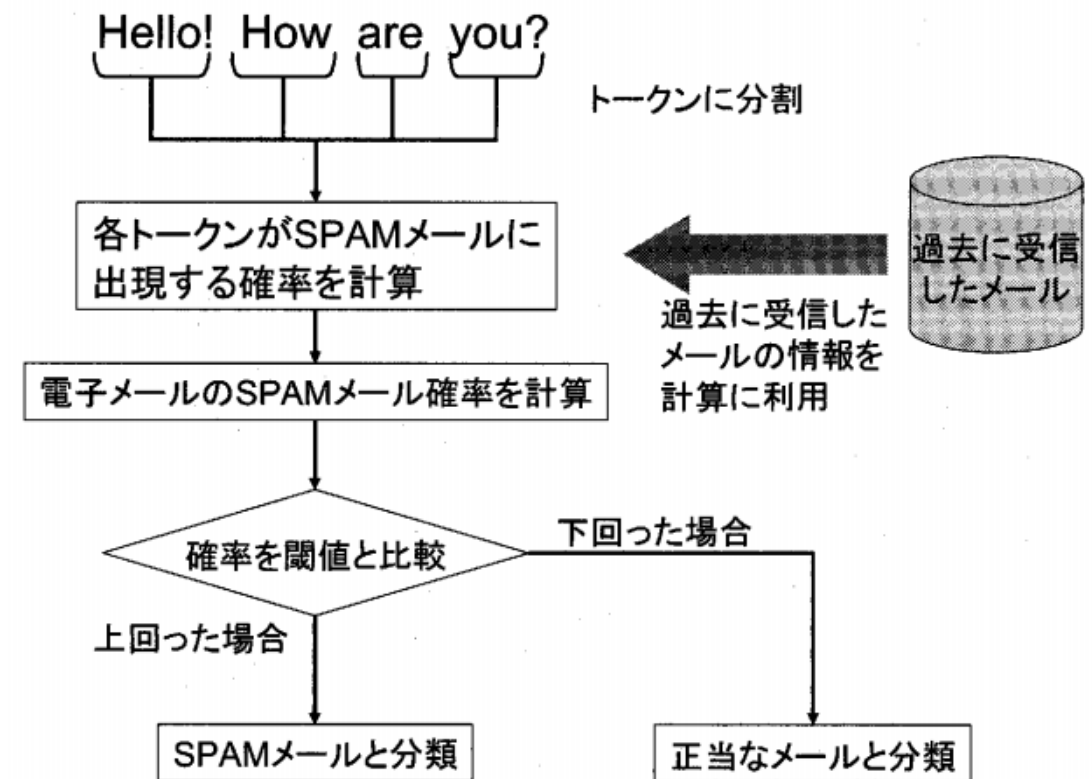
アルゴリズム

- 全学習データについて誤差関数 $E(\mathbf{W})$ とその勾配を計算する
- 上記の式で全ての重みとバイアスを更新する
- 誤差が十分小さくなるまで1, 2を繰り返す



事例 | スпамメールフィルタ (Gmail, Yahoo!メールなど)

- スпамメールの分類には教師あり学習である**単純ベイズ分類器**や**ニューラルネットワーク**が主に使われてきた
- メールの文章を単語ごとに分析し、スパムメールによく使われる単語かどうかを元にメール自体がスパムであるかどうかを判断
 - そのせいで普通のメールで文章内に「突然ですが」や「登録お願いします」などを入れたがためにスパム扱いされることもある
- ユーザがスパム扱いしたかどうかを受けて逐次学習している

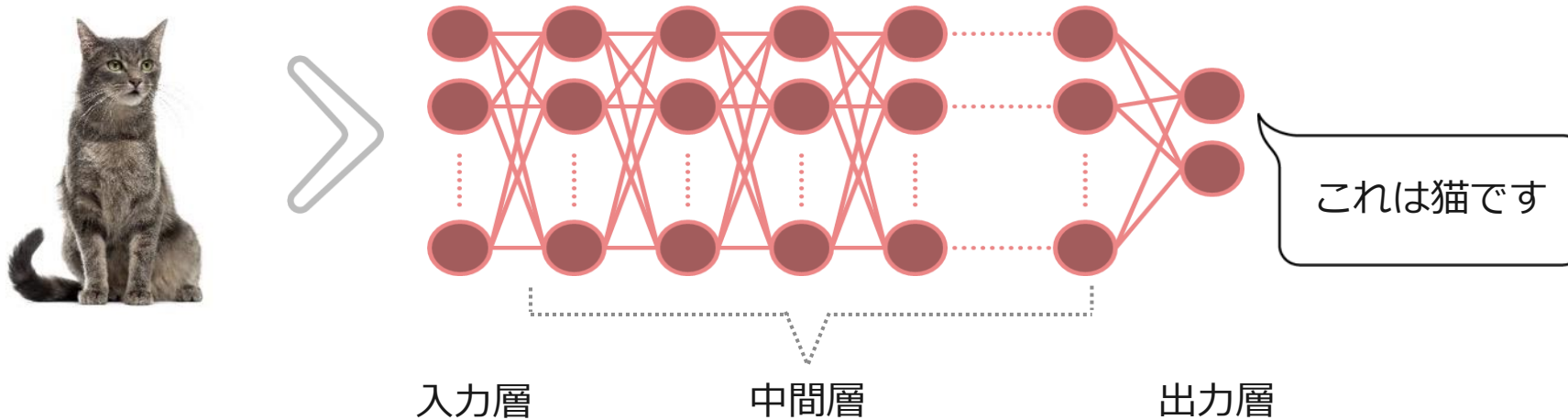


▶ [SPAMメールフィルタリング：ベイジアンフィルタの解説\(<特集>情報のフィルタリング\)](#)

2. ディープラーニングの概要

ディープニューラルネットワーク

- Deep Neural Network (DNN、多層ニューラルネットワーク)
 - 入力層を除いて3層以上からなるニューラルネットワーク
- ディープラーニング (深層学習)
 - 主に多層ニューラルネットワークを用いた機械学習
- DNNの成果
 - 画像認識分野で顕著
 - 自然言語処理や音声認識等への適用研究も進んでいる



ディープラーニングと従来の機械学習の比較

従来の機械学習



ディープラーニング

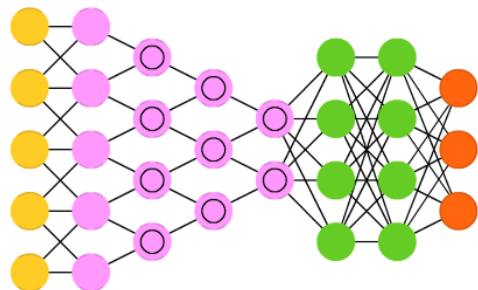


ニューラルネットワークの種類（再掲）

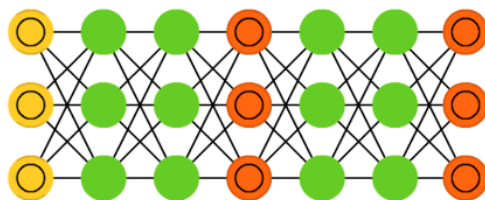
- ・ 教師あり学習の機械学習（自己符号化器など一部は教師なし）
- ・ 左から右に一方通行で計算を行っていく順伝播型
- ・ 逆方向への経路を持つ再帰型

順伝播型（フィードフォワード）

畳み込みニューラルネット

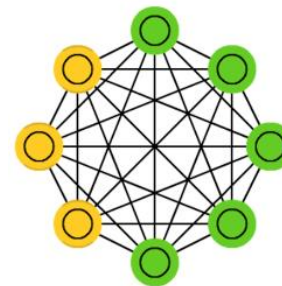


Generative Adversarial Networks



再帰型（リカレント）

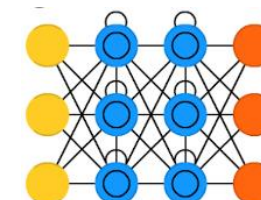
ボルツマンマシン



自己符号化器



LSTMネットワーク



▶ 画像の出典 (2017.09.04) <http://www.asimovinstitute.org/neural-network-zoo/>

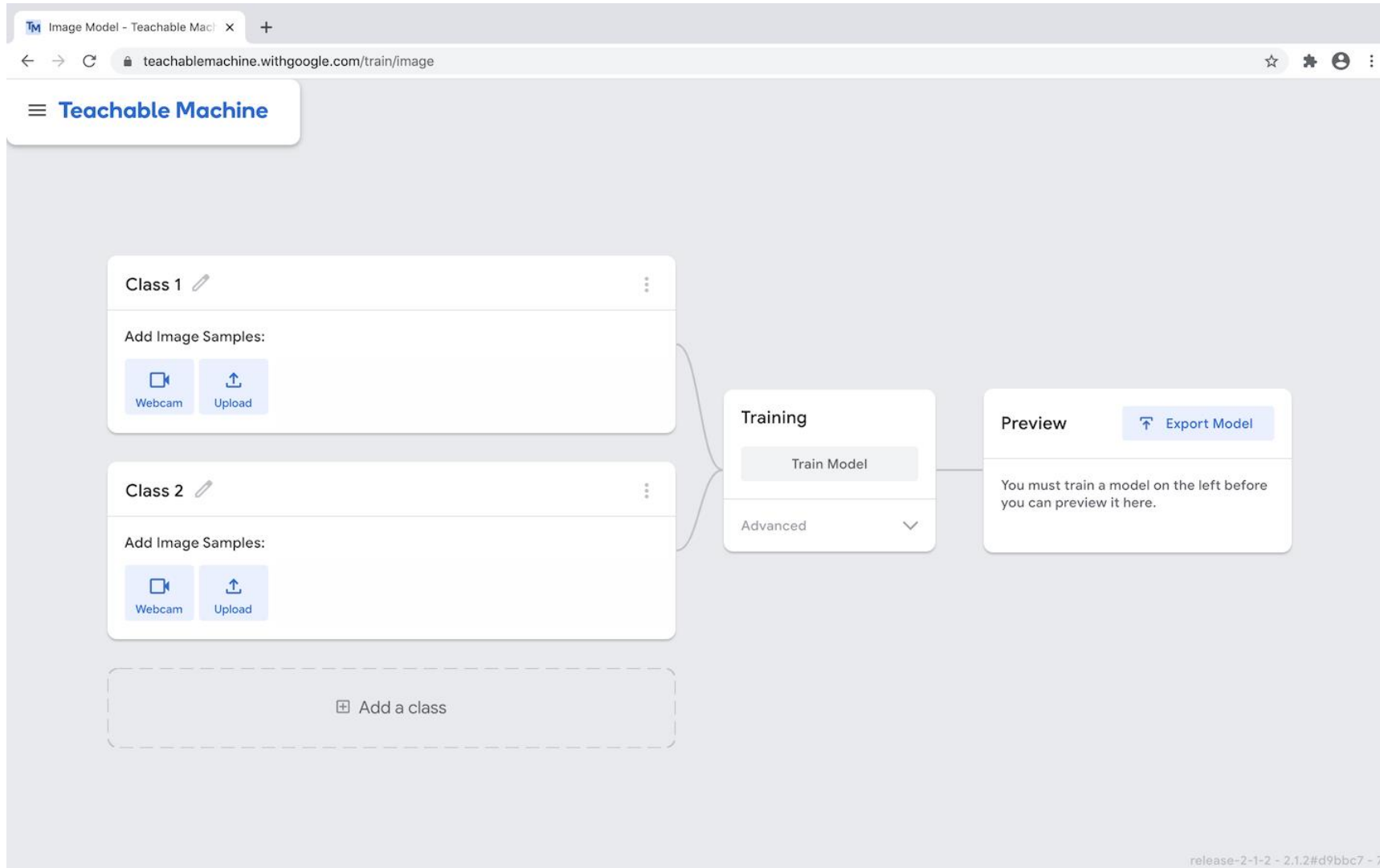
参考 | ディープラーニングとGPU

- ディープラーニング研究・活用が大きく進展した要因としてGPUの低価格化が寄与
 - ハイエンドCPUの10倍以上の性能を持つGPUがほぼ同価格で手に入る時代に
 - 1週間かかる計算が1日で終わるように、**1年かかる計算もGPU4枚なら1週間程度**で終わる
- 科学計算・研究用途を主に販売されているTeslaと一般向けのGeForceがある
 - 中でもGeForce最高クラスの「**TITAN RTX**」は価格は30万円程度とTesla V100の1/4程度、性能も近いのでディープラーニング用GPUとして広く使われている（2、4枚刺しが多い）
 - 「**RTX 2080 Ti**」はメモリサイズこそ「TITAN RTX」の半分以下であるもののそれ以外のスペックは近く、価格は「TITAN X」の半額程度でコストパフォーマンスがよい（2020年7月現在）
 - ただしGeForceシリーズは一般向けモデルのためTeslaよりも信頼性は劣る
 - モデルによってNvidiaの数理計算高速化ドライバ「CUDA」や「cuDNN」のサポートするバージョンや対応の可否が異なる



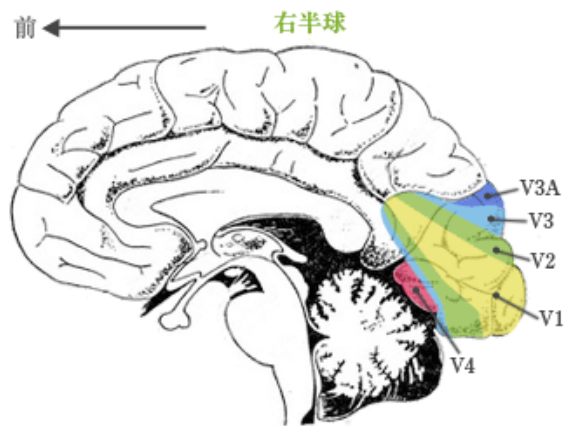
▶ Nvidia社製GPU Tesla
ハイエンドCPUの10倍以上の計算
性能を持つ。コア数は100倍以上。

ディープラーニングによる画像認識のデモ

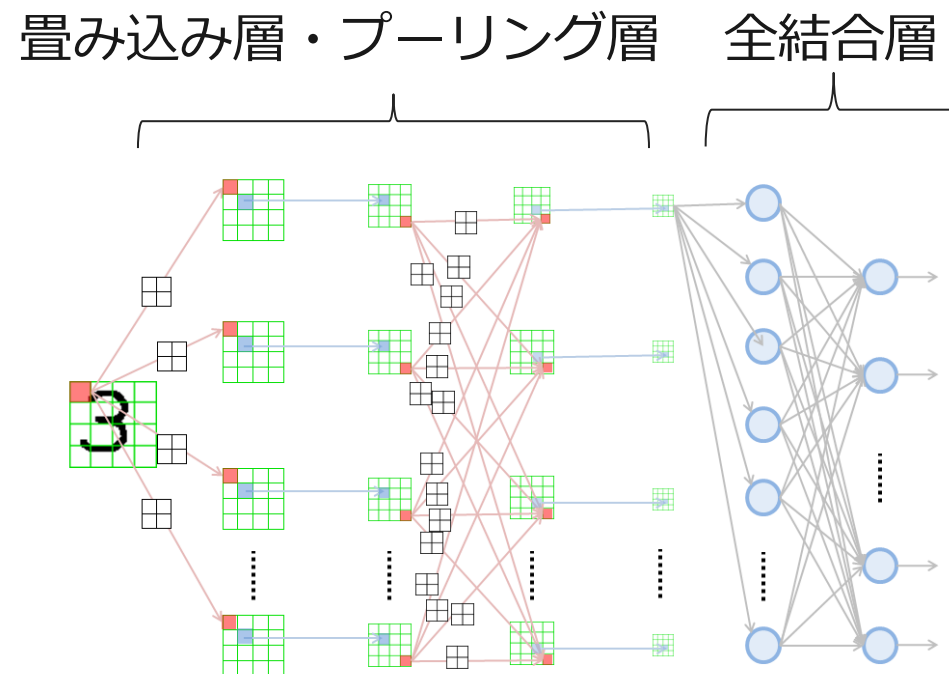


畳み込みニューラルネット

- Convolutional Neural Network (CNN)
- 畳み込み層・プーリング層・全結合層からなる順伝播型ネットワーク
- 動物の脳の視覚野の構造を模している
 - 階層的な構造を持つ（1次視覚野、2次視覚野、．．．）
- 画像認識に適している
 - 自然言語処理、数値分析、音声認識などでも一部使用



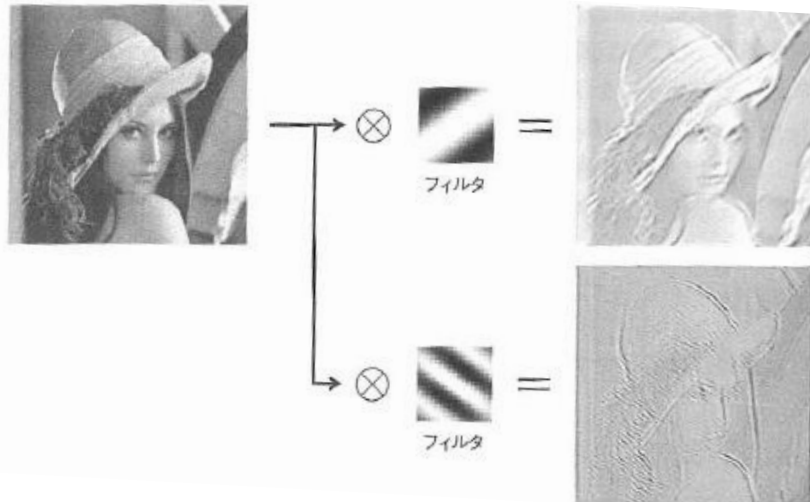
▶ <http://whatsdesign.arrow.jp/basic08.html>



畳み込み層とプーリング層

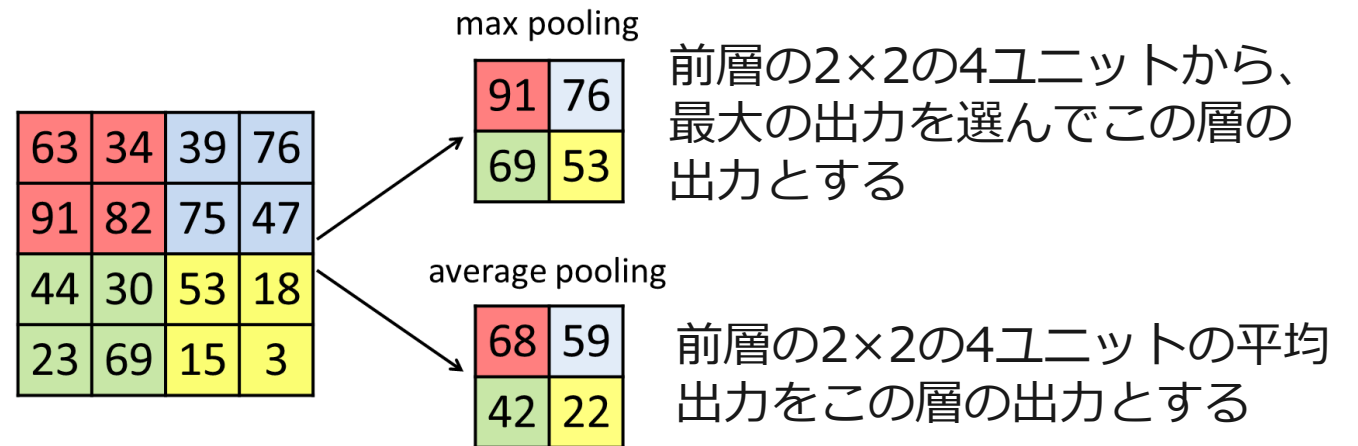
- CNNでは畳み込み層を1～2回程度繰り返した後にプーリング層が続くことが多い
 - **畳み込み層**：データの中から、**フィルタのパターンと類似したパターンを検出**
 - **プーリング層**：畳み込み層が抽出した**パターンを圧縮**し位置的なズレを緩和
- 畳み込み層とプーリング層は全結合でないことが普通
- 畳み込みとプーリングを繰り返すことにより、データの特徴を抽出

畳み込み処理のイメージ



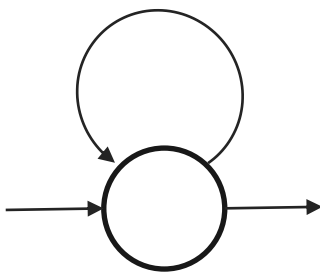
▶ 講談社 MLPシリーズ「深層学習」より

プーリング処理のイメージ

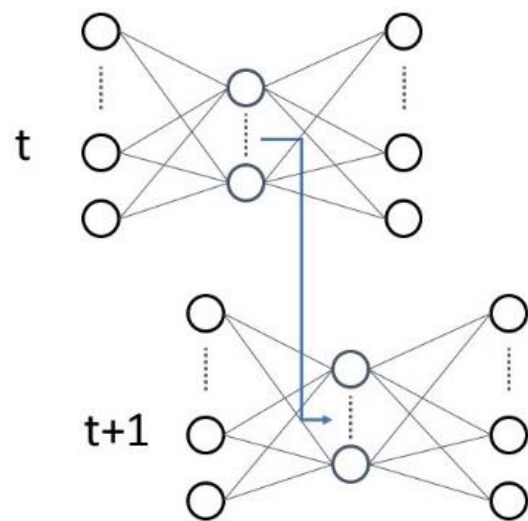


リカレントニューラルネット

- Recurrent Neural Network (RNN)
- 自分自身に入力する再帰的な結合構造（閉路）を持つニューラルネットの総称
- 順序に意味がある音声、言語、動画などの時系列・順序データを扱うのに適する
- 文が途中まで与えられた時、次に来る単語を予測
 - 「私は家でご飯を」の後に「食べる」が来ることを予測
- 課題：長期に渡ってのデータ参照がしにくい

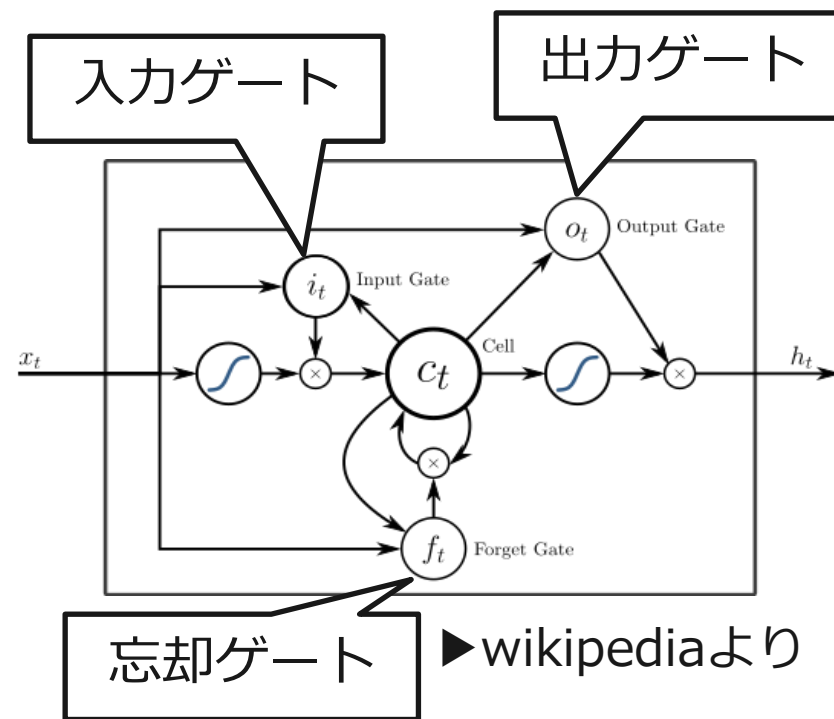


前の層の入力に加えて自身の
前回の出力を入力とする



時間軸で展開して書くと
左記のようになる

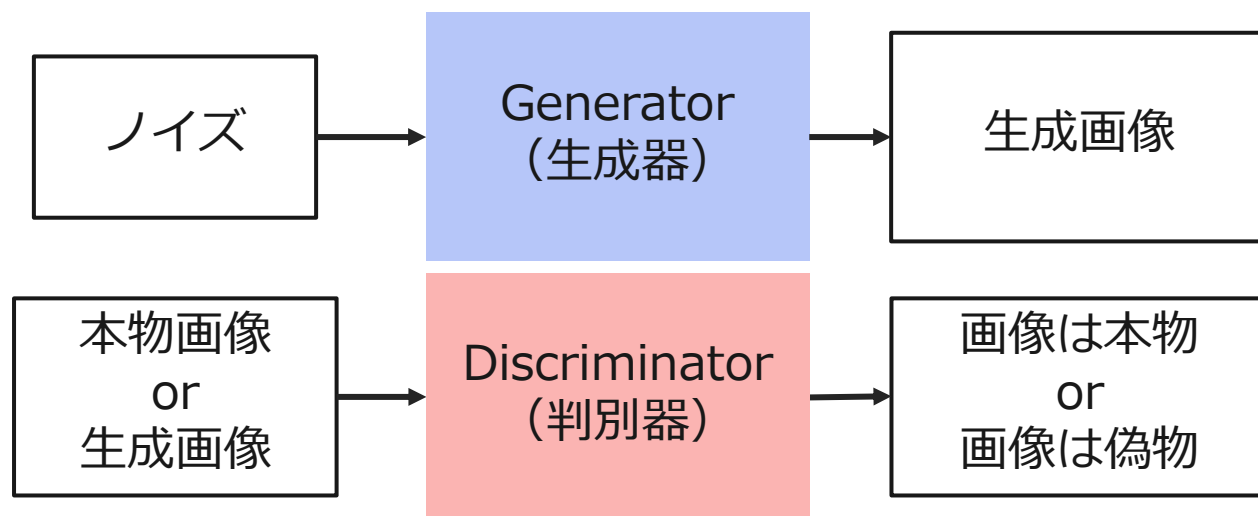
- Long Short-Term Memory (LSTM)
- RNNの一種で、より長期的な過去の履歴を扱うことができるように改良されたモデル
- 基本的な構造は元のRNNと同じだが、中間層のユニットをメモリユニットで置き換えた構造を持つ
 - 入力ゲート、出力ゲート、忘却ゲートの3つ
- 忘却ゲートでメモリをどれくらい保持するか調整
- 音声認識、自然言語処理のディープラーニングとして一部は実用レベルでも使用



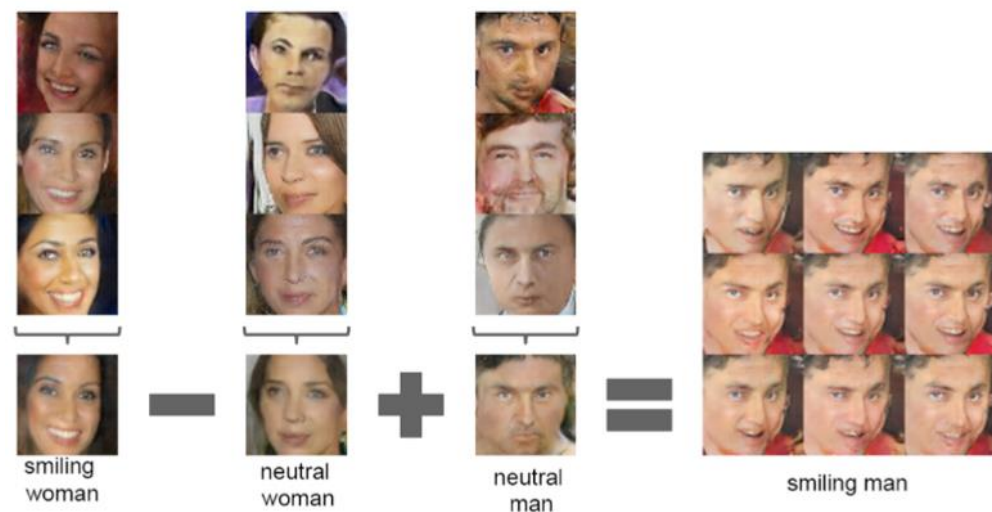
敵対的生成ネットワーク

- Generative Adversarial Networks (GANs)
- 2つのネットワークを敵対的に（競わせることで）訓練する生成モデル
 - Generator（生成器）は訓練データを学習し、訓練データと似た特徴を持つデータを生成
 - Discriminator（識別器）は訓練データとGeneratorが生成したデータを識別

GANsの仕組み

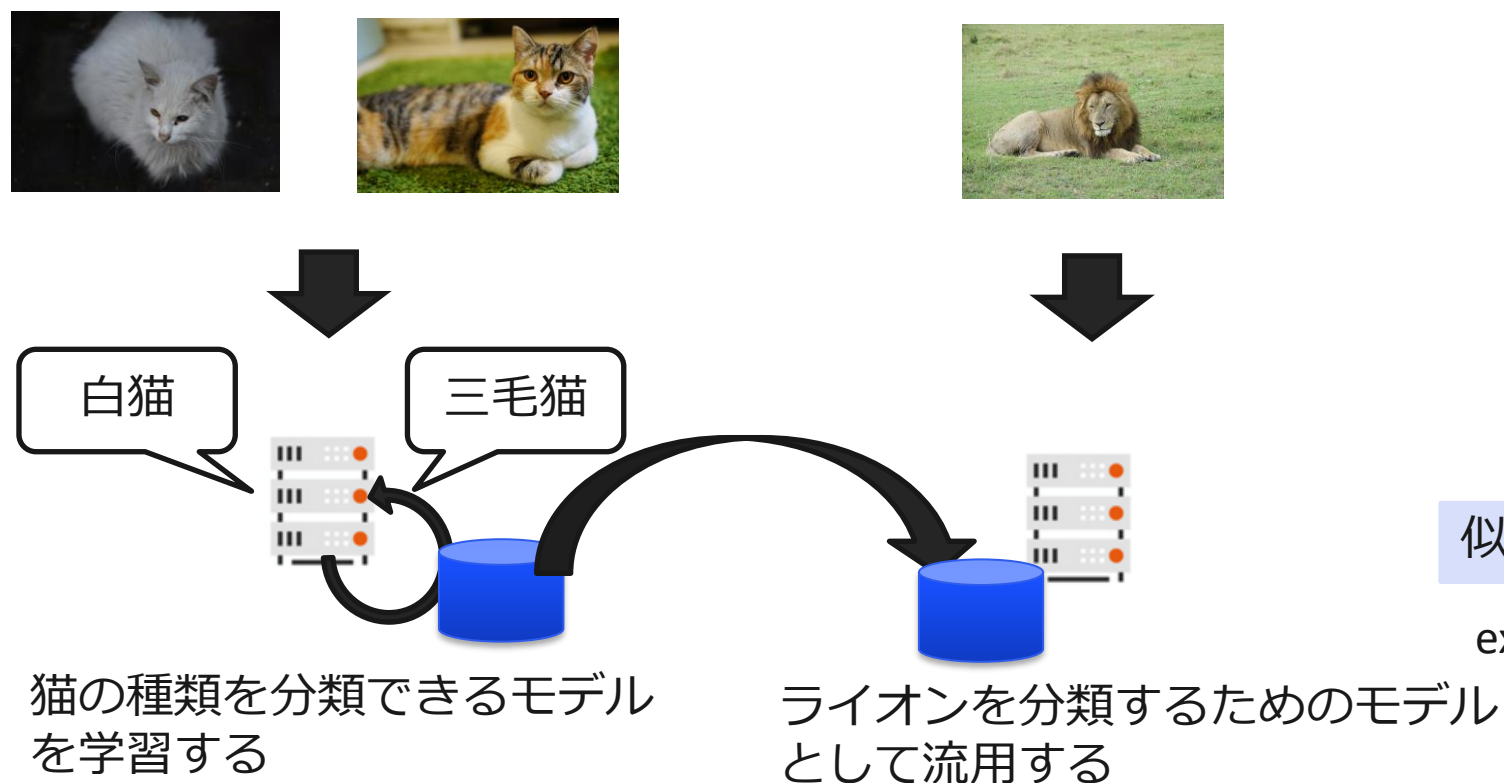


CNNによるGANsの実装例



► Radford and Chintala 2015

- 転移学習：機械学習において、ある領域で学習させたモデルを別の領域に適用（流用）させる技術



転移学習の利点

- (適切な領域や手法を選択すれば)
- ・一から学習しなおすよりも収束が早い
 - ・データが少なくても精度が出せる
 - ・共通する特徴をうまく使うよう学習させることにより汎化性能を高められる

似た領域（ドメイン）で適用するのがコツ

ex) 車の種類を増やす、有名人を増やすetc...

3. 活用事例

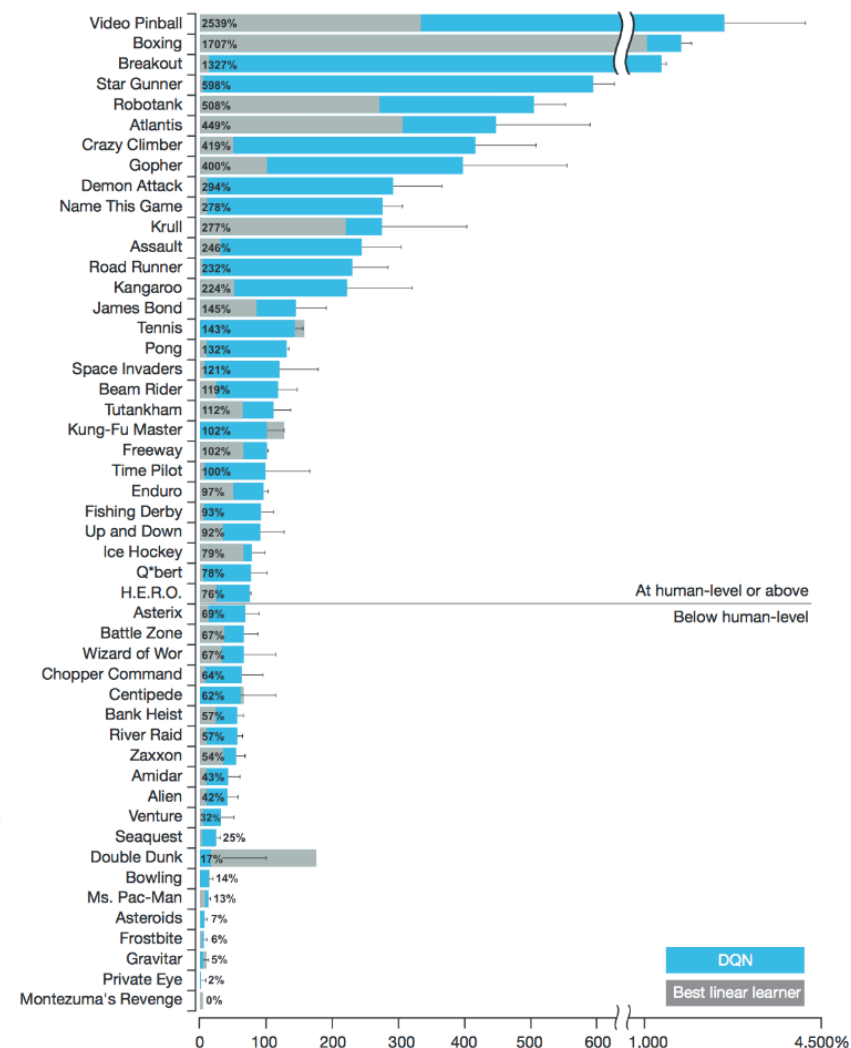
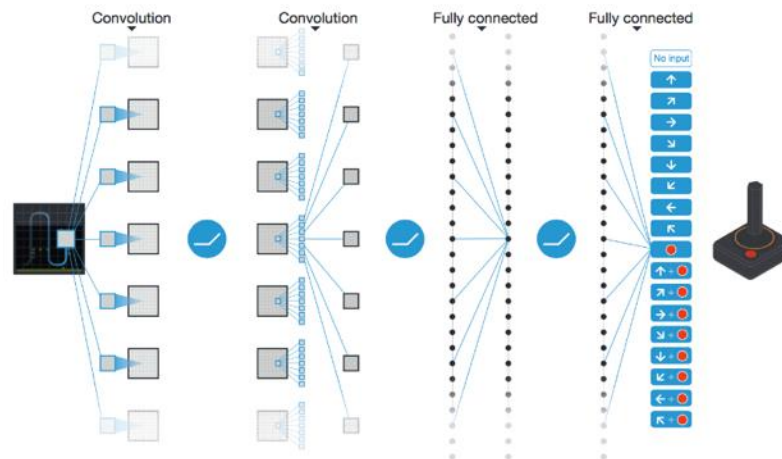
深層学習によりゲームでハイスコアを記録

Human-level control through deep reinforcement learning

Volodymyr Mnih^{1*}, Koray Kavukcuoglu^{1*}, David Silver^{1*}, Andrei A. Rusu¹, Joel Veness¹, Marc G. Bellemare¹, Alex Graves¹, Martin Riedmiller¹, Andreas K. Fiedjeland¹, Georg Ostrovski¹, Stig Petersen¹, Charles Beattie¹, Amir Sadik¹, Ioannis Antonoglou¹, Helen King¹, Dhharshan Kumaran¹, Daan Wierstra¹, Shane Legg¹ & Demis Hassabis¹

畳み込みニューラルネット（CNN）と強化学習を組み合わせたDeep Q-Netowrk（DQN）を用いて、Atari 2600の46ゲーム中29ゲームで人間と同等以上（一部は圧倒）のスコアを記録

ゲーム画面の画像を入力として、ゲームコントローラーのコマンドを出力とするCNNを設計



► Mnih et al. 2015

DQNによるゲームプレイ動画



Mastering the game of Go with deep neural networks and tree search

David Silver^{1*}, Aja Huang^{1*}, Chris J. Maddison¹, Arthur Guez¹, Laurent Sifre¹, George van den Driessche¹, Julian Schrittwieser¹, Ioannis Antonoglou¹, Veda Panneershelvam¹, Marc Lanctot¹, Sander Dieleman¹, Dominik Grewe¹, John Nham², Nal Kalchbrenner¹, Ilya Sutskever², Timothy Lillicrap¹, Madeleine Leach¹, Koray Kavukcuoglu¹, Thore Graepel¹ & Demis Hassabis¹

深層学習、強化学習、モンテカルロ木探索を
組み合わせて作られた囲碁ソフトAlphaGoが
世界トップ棋士のイ・セドル九段に勝利

2016年末

AlphaGo Masterが世界最強に

2017年10月

AlphaGo ZeroがMasterに圧勝

2017年12月

AlphaZeroがチェス・囲碁・将棋で世界最強に

世界トップ棋士のイ・セドルにGoogle AIが挑んだ五番勝負、4勝1敗でAIが勝利

2016年3月16日 by *Jon Russell*

58

List

15

f シェア

ツイート

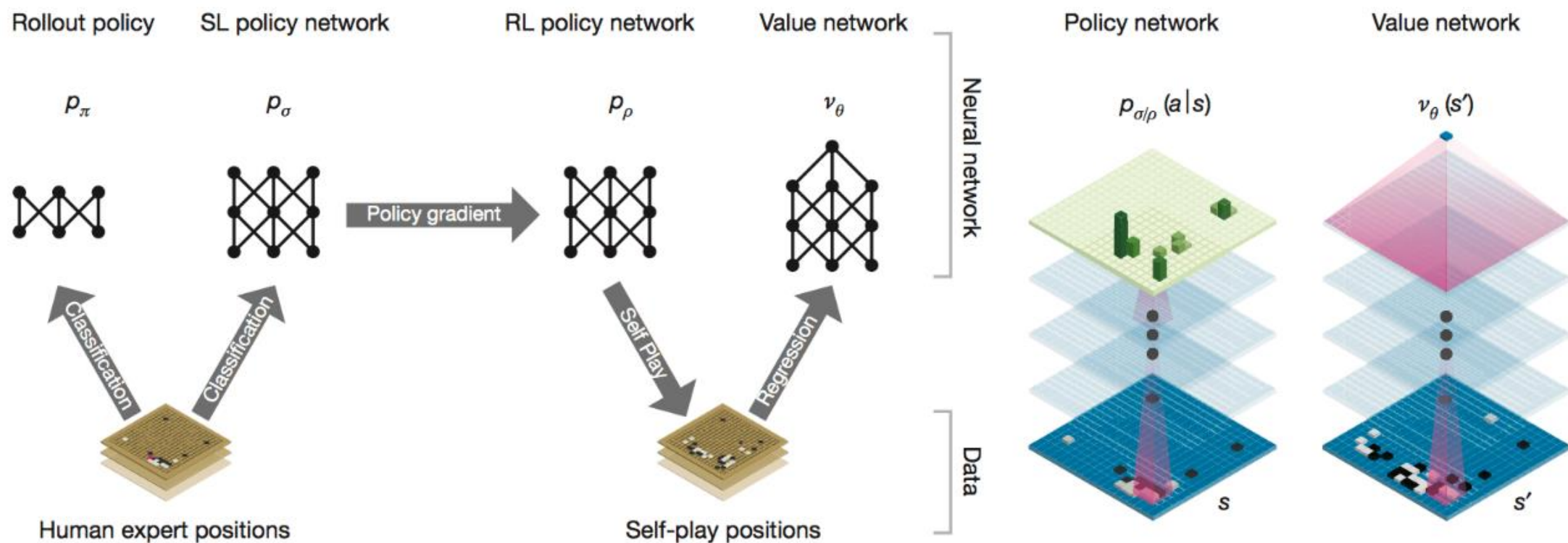
BI はてな



▶ TechCrunchニュース（2017/09/01に取得）
<http://jp.techcrunch.com/2016/03/16/20160315google-ai-beats-go-world-champion-again-to-complete-historic-4-1-series-victory/>

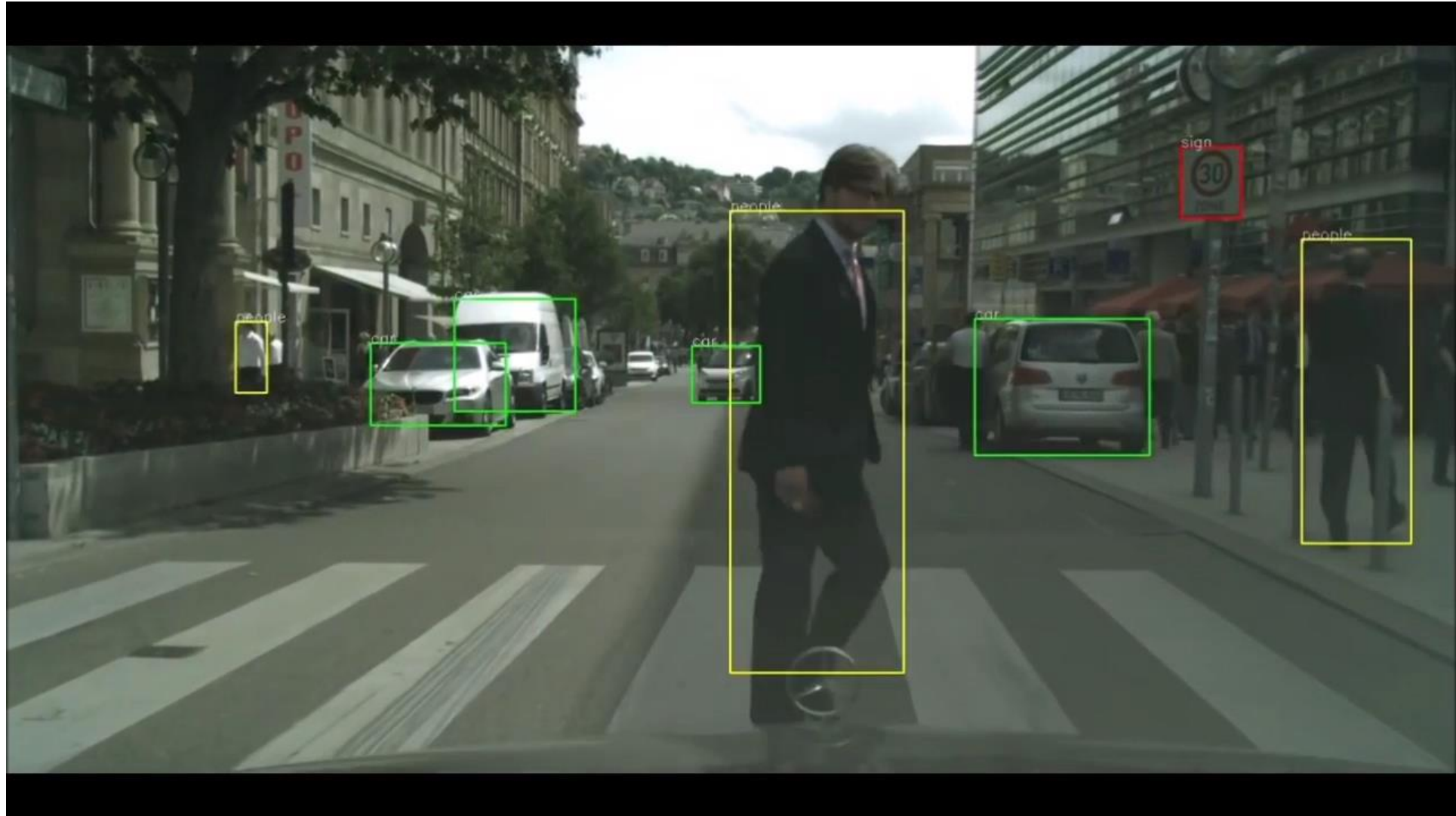
深層学習を用いた囲碁ソフト AlphaGo

Policy network (打つ手を決める) と
Value network (盤面から勝率を計算) をそれぞれCNNで作成



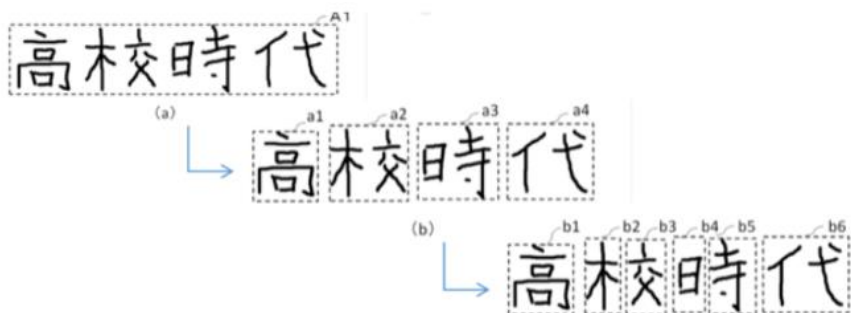
► Silver et al. 2016

Nvidia車載画像認識のデモ



日本語の文字認識 (OCR)

- 日本語は文字の種類が多いため、OCRの精度向上が難しい
 - アルファベットは52文字、日本語は常用漢字、人名漢字含め3000字以上
- Tegaki
 - Cogent Labsが開発し、APIとして提供している
 - 5人が書いた文章全体で識字率99.2%
- DX Suite (Intelligent OCR)
 - AI insideが開発し、クラウドサービスとして提供している
 - 深層学習と自然言語処理を用いて精度を向上 (特許取得済み)
 - ・ 漢字の辺やつくりを別個に認識し、整合性をチェック
 - ・ 文字ごとの認識だけでなく、前後の文字との繋がりから単語として妥当かを判断



▶ <https://inside.ai/company/patent.html> (2017.07.08)

Google データセンターの省エネ化

■ 目的

- データセンターの冷却設備の設定の最適化

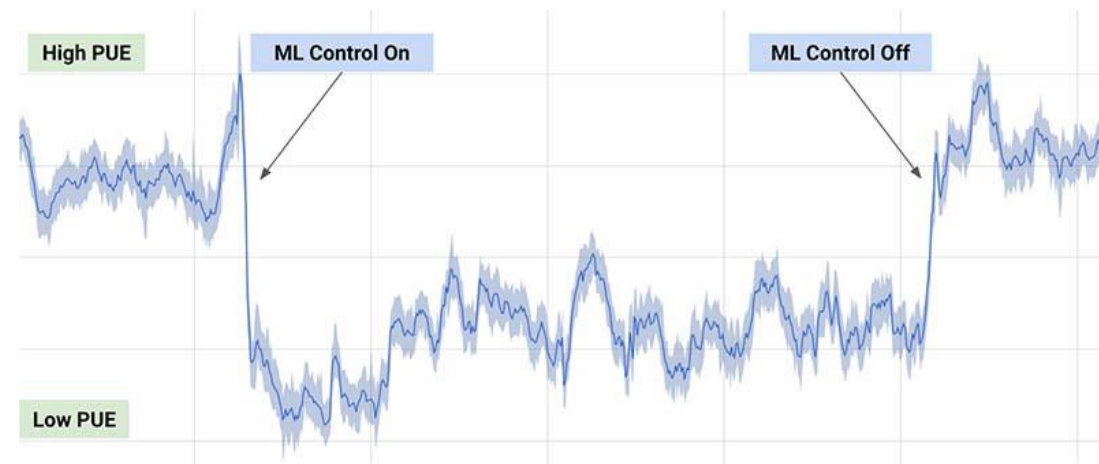
■ 手法

- DeepMindの深層強化学習技術を応用
- データセンター内に数千のセンサーを配置
温度、電源、ポンプ速度、各種設定値などのデータを収集
- 上記データと電力消費量、電力使用効率の関係を学習
- 学習結果より冷却設備の効率的な運用シナリオを予測・運用

■ 結果

- 電力消費量の40%削減に成功

AIに運用を任せた期間の省エネ化に成功



▶ <http://itpro.nikkeibp.co.jp/atcl/news/16/072102162>

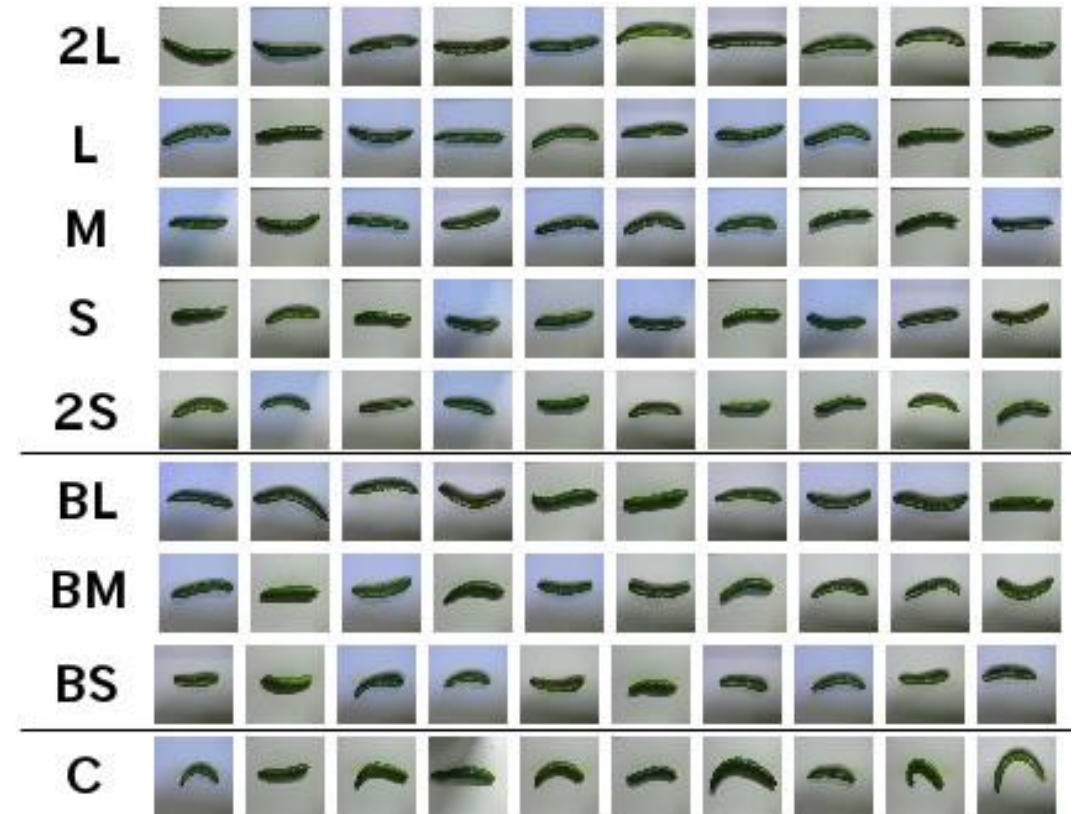
TensorFlowによるきゅうりの仕分け

■ 背景

- きゅうり農家である母親の仕分け作業は1日8時間の長時間労働
- 元組み込みエンジニアのスキルを活かし、TensorFlowで仕分け機を自作

■ 目的

- きゅうりを9つの等級に自動分類



▶ <http://workpiles.com/2016/02/TensorFlow-cnn-cucumber>

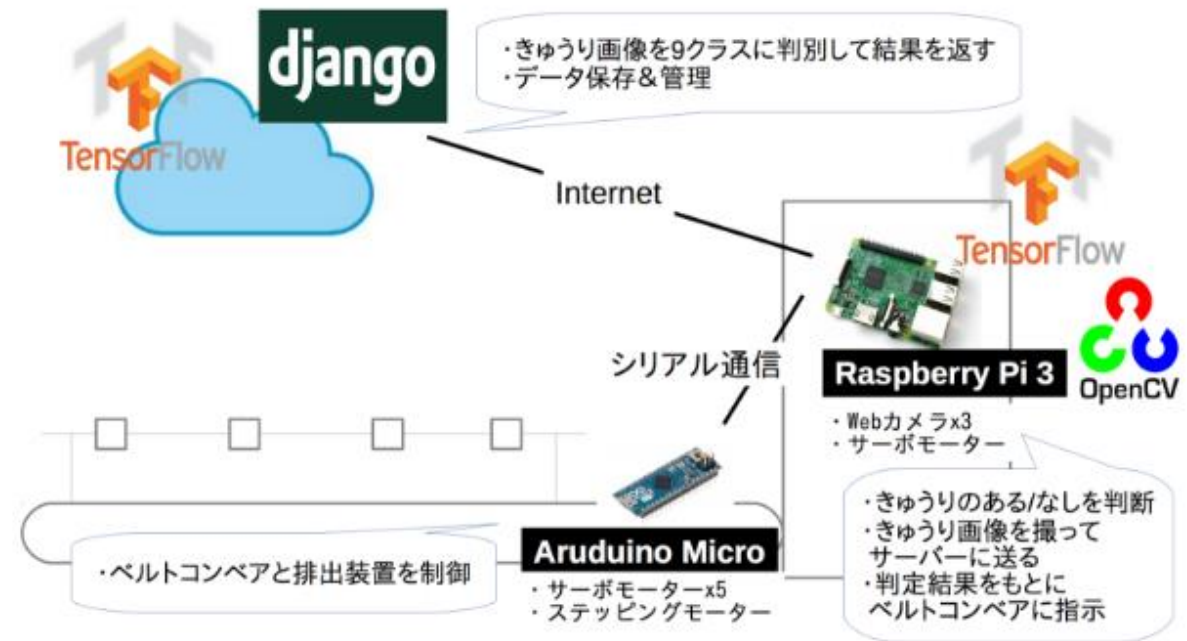
TensorFlowによるきゅうりの仕分け



► https://cloudplatform-jp.googleblog.com/2016/08/TensorFlow_5.html

TensorFlowによるきゅうりの仕分け

- 入力データ
 - 80x80ピクセルの低解像度画像
- 学習
 - 2ヶ月かけて7000枚の学習データ
1500枚のテストデータを用意
 - 学習完了まで自宅PCで2・3日
- 結果
 - テストデータでの正解率は95%
 - 本番データでは70%程度
→ 過学習が発生



▶ https://cloudplatform-jp.googleblog.com/2016/08/TensorFlow_5.html

ニコニコ動画のコメント解析におけるディープラーニング

■ 応用: コメント監視

- 誹謗中傷や殺人予告などの不適切コメントの削除
- 人が目視で監視
- 量が膨大でタイムラグも大きい

■ コメントの識別器

- コメントの黒・白・グレーを判別

■ 結果

- 全体の7割が監視対象外に
- 目視による監視を最低限に抑えることに成功

2015年紅白歌合戦での
小林幸子氏とのコラボで
もコメントが話題に

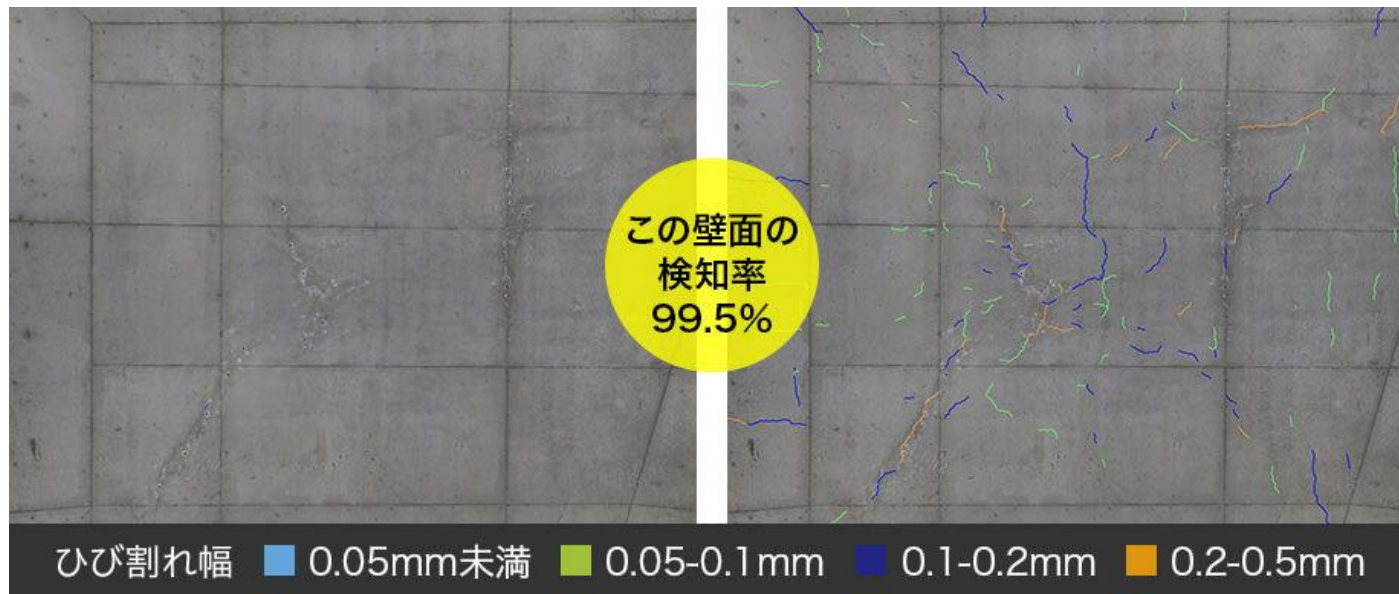
黒：誰が見てもアウト
白：人が見なくてもOK
グレー：人による判断が必要

▶ <http://ainow.ai/2017/04/20/111224>

▶ <https://news.mynavi.jp/article/deeplearning-6>

ヒビ割れ検出（キヤノン）

- コンクリート壁面のヒビ割れをディープラーニングの画像認識を用いて検出
- 人が720分を要する作業を、AIで90分で点検結果を作成できた例も
- 人間だと作業者によってばらつく点検結果を、AIでは常に一定の品質で作成可能



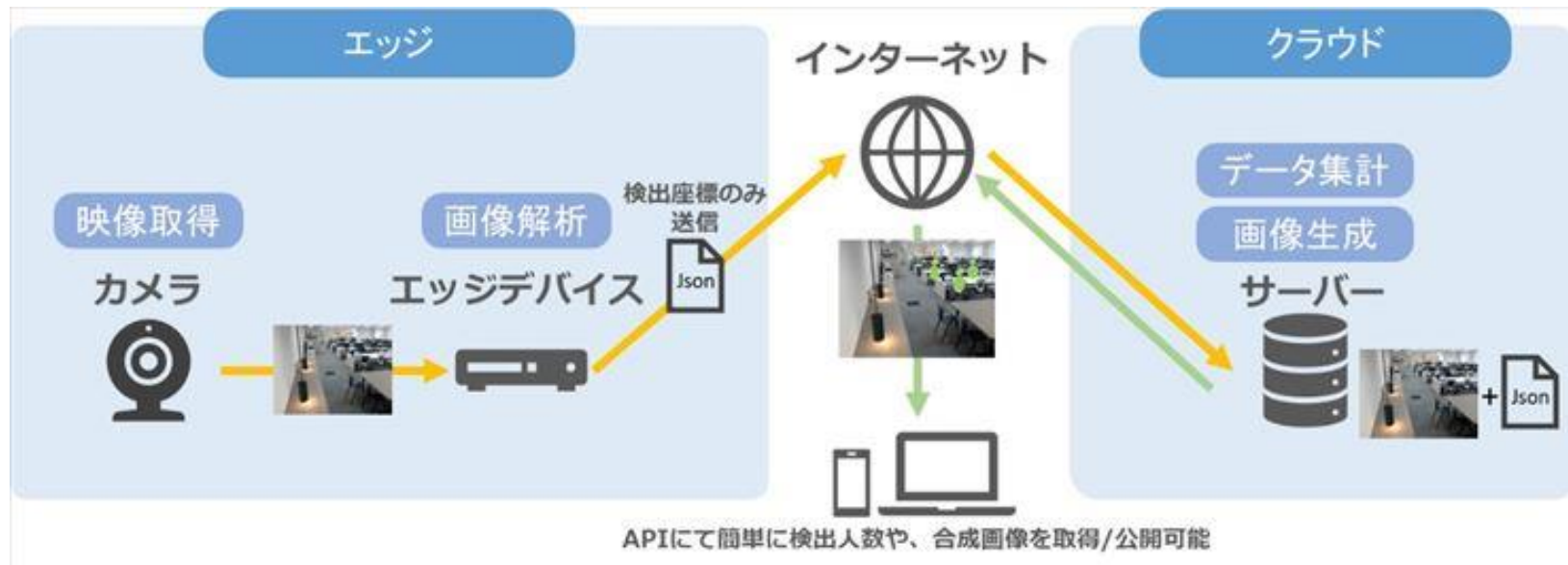
▶ <https://global.canon/ja/technology/crack2019.html>



撮影プロセスもキヤノン側から提供することで細かいひびを検出できる高精度の画像を撮影（ドローンを使う場合もある）

AIカメラ活用（空席状況の見える化）

- ディープラーニング搭載のAIカメラを用いて店内の人物を検出し空席状況の見える化
- AIカメラ活用はこれ以外でも開発が進んでおり、マーケティングや不審者検出、放置自転車の検出などの応用例も



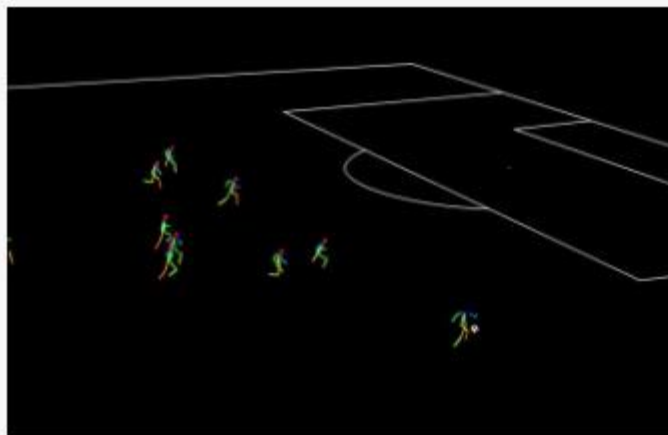
▶ <https://prtmes.jp/main/html/rd/p/000000031.000018045.html>

スポーツ分析(PFN)

- サッカーの中継映像から各選手の姿勢、向きを画像認識により検出
- さらに試合の状況（フリーキック中、カウンターされている等）も分類する



▶ <https://preferred.jp/ja/projects/sports-analytics/>



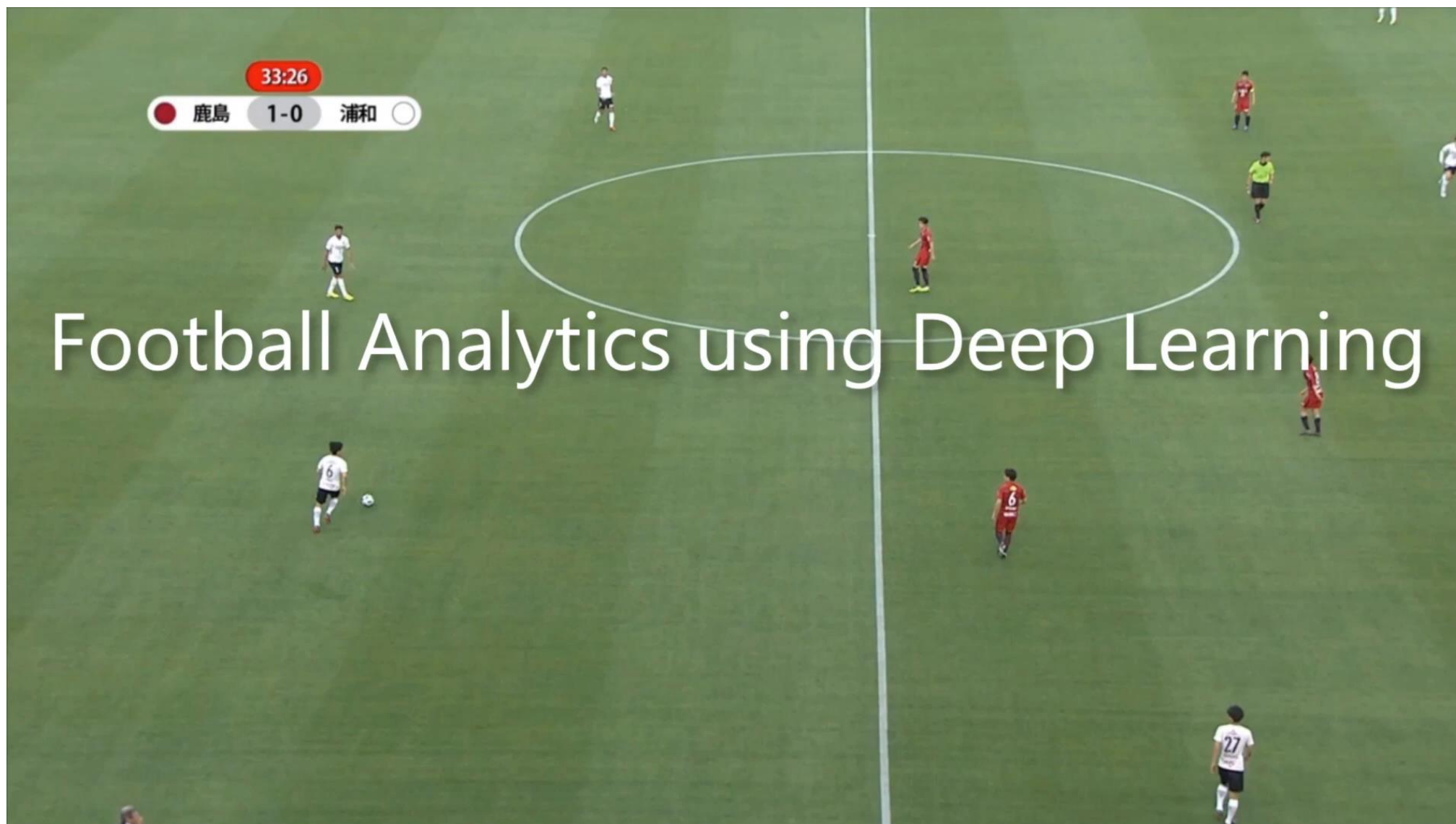
選手の姿勢推定（特許出願済）



姿勢推定にもとづく選手の向き検出（特許出願済）



サッカーシーンの分類（特許出願済）

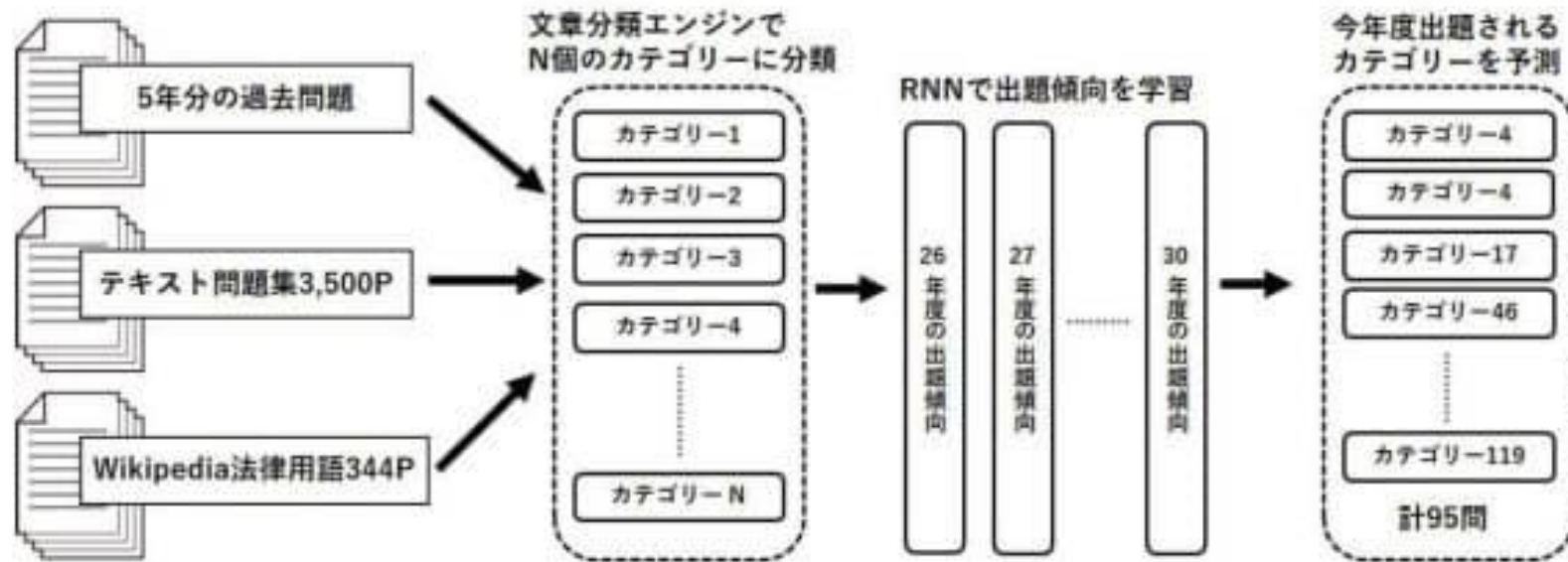


▶ <https://preferred.jp/ja/projects/sports-analytics/>

© MindTech inc.

試験問題予測AI

- 司法試験予備試験の出題傾向の予測を行うAI
- ディープラーニングで過去の出題傾向を学習しており、次の年の傾向予測ができる
- 6割程度の予測精度が出せている（司法試験の合格基準は6割）



▶ <https://www.nikkei.com/article/DGXMZO44236150W9A420C1000000/>

4. ライブラリとフレームワーク

深層学習ライブラリ・フレームワーク

■ TensorFlow

- Googleが開発した深層学習ライブラリ

■ Keras

- TensorFlow等のライブラリを容易に利用できるラッパーライブラリ
- 現在はほぼTensorFlow用

■ Chainer

- Preferred Networksが開発していた深層学習ライブラリ

■ PyTorch

- Facebookなどが主導するフレームワーク
- 使用感はChainerに近い

クラウドディープラーニング用インフラ

■ AWS GPUインスタンス

- EC2のインスタンスの一種として提供されるP2インスタンス
- GPUは1,8,16から選べる
- 金額は一時間あたり100円～1500円程度(1日2500～36000円)

■ IBM Bluemix Infrastructure

- 元々SoftLayerという名前で提供されていた
- ベアメタルサーバ（物理サーバ）のみの提供
- 月額約80000円～

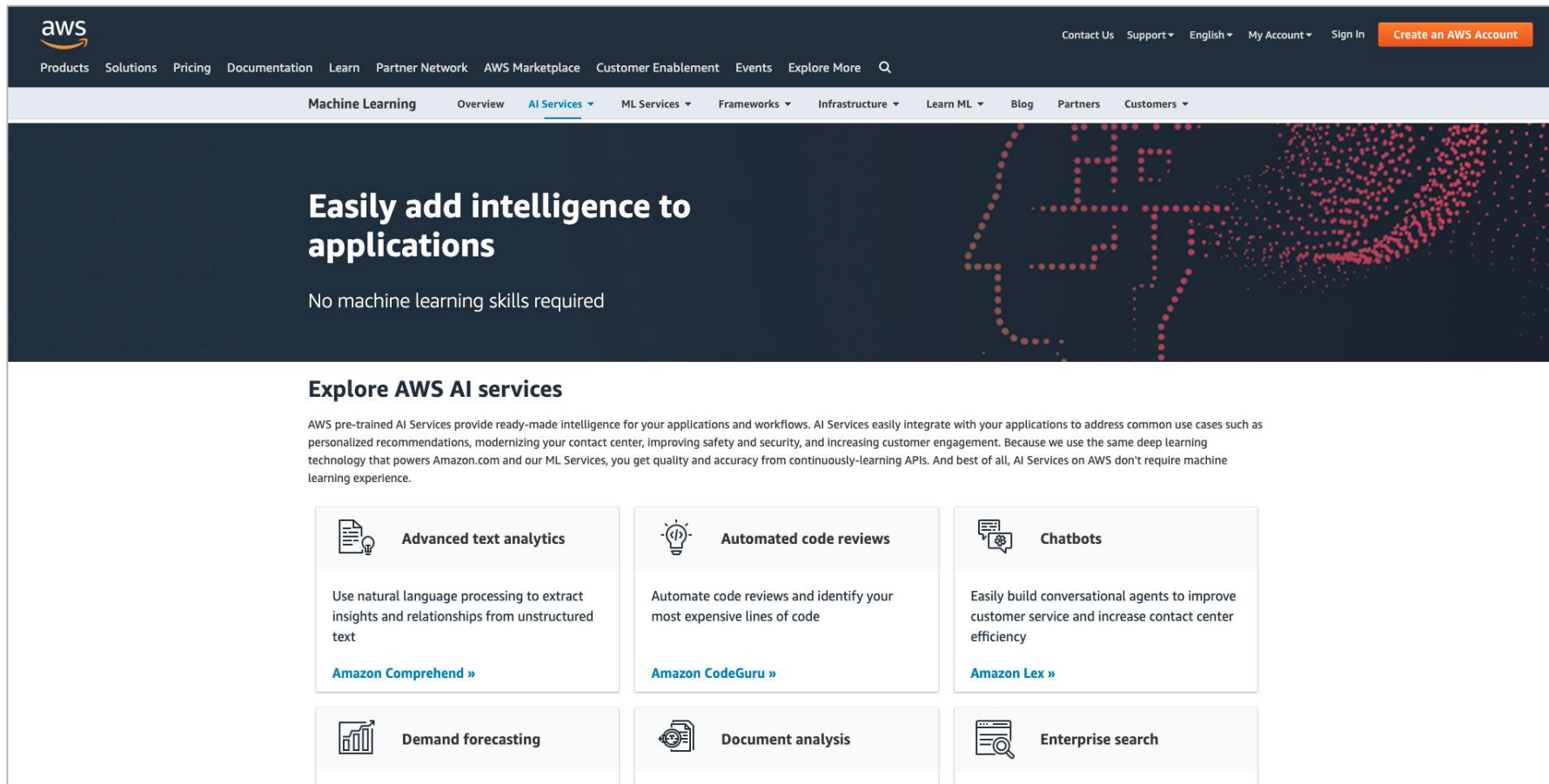
■ Azure N シリーズ

- GPUは1,2,4個から選べる
- 一時間あたり110円～500円程度

■ その他GCPやさくらインターネットなどからも提供されている

ディープラーニング系主要プラットフォーム（1/4）

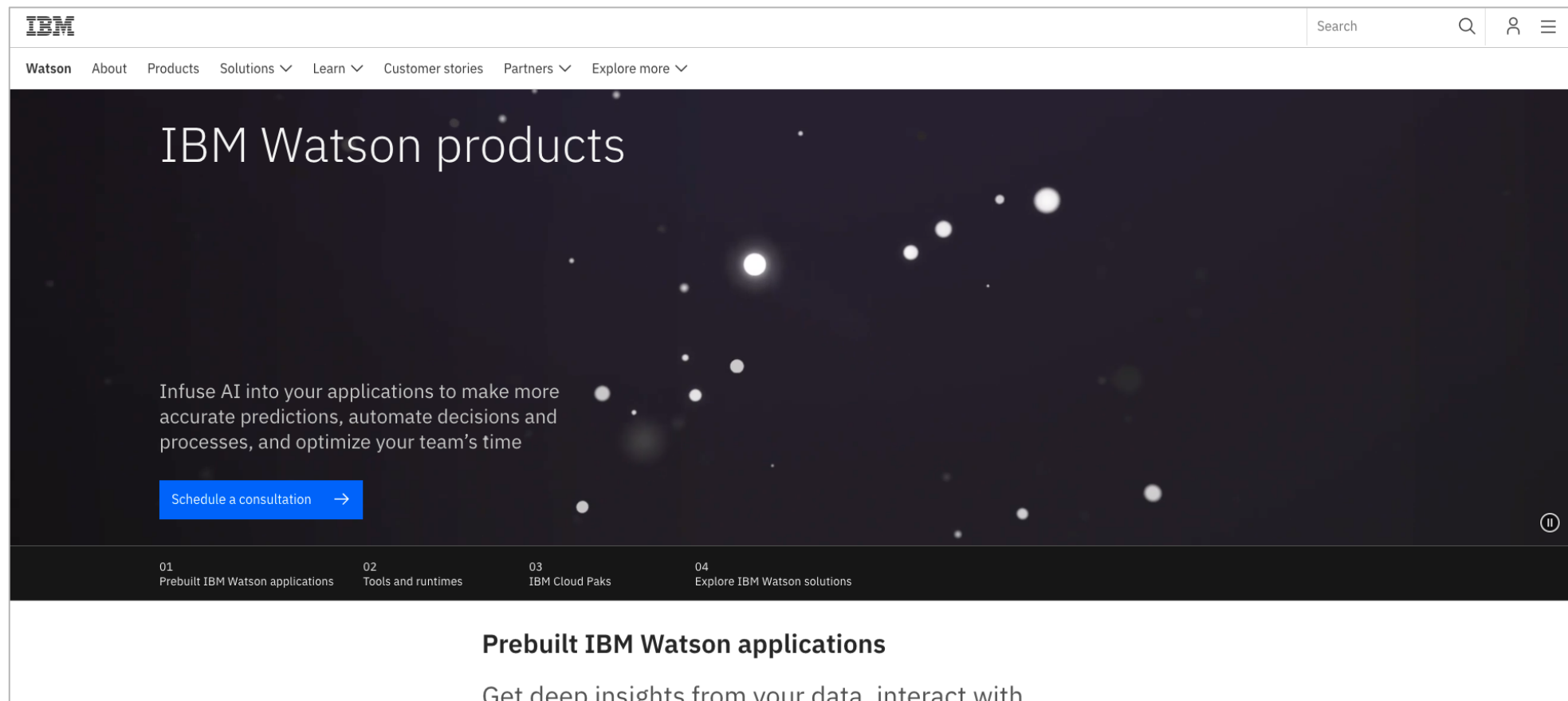
- AWS AI Services (<https://aws.amazon.com/machine-learning/ai-services/>)
画像認識サービス、チャットボット作成サービス、自然言語処理サービス、
その他、需要予測サービス等も提供



ディープラーニング系主要プラットフォーム (2/4)

■ IBM Watson

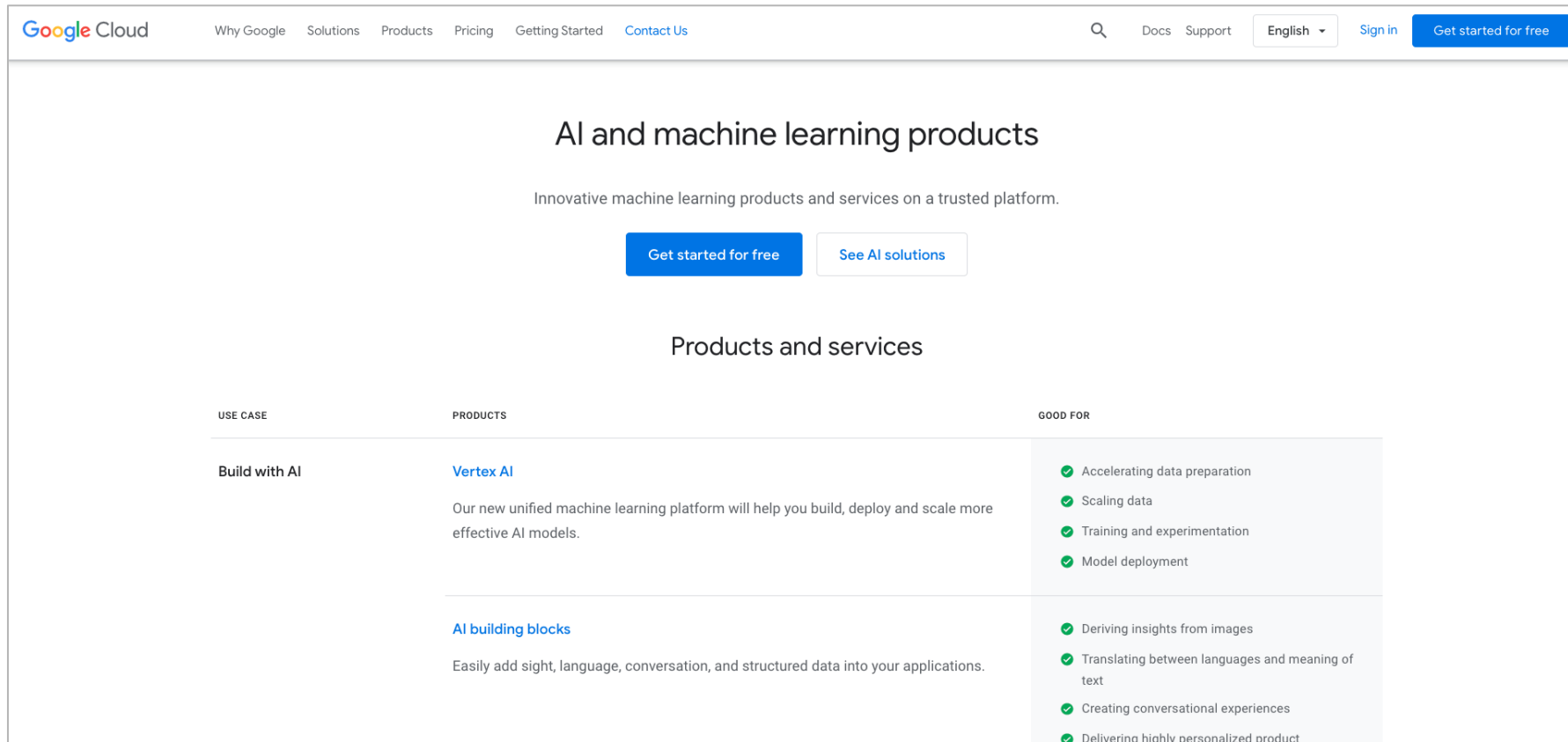
- <https://www.ibm.com/watson/products-services>
- Watsonを有名にした自然言語処理、音声認識に強い
- 日本語対応はソフトバンクの協力の元なので語彙力がある



ディープラーニング系主要プラットフォーム（3/4）

■ Google Cloud ML

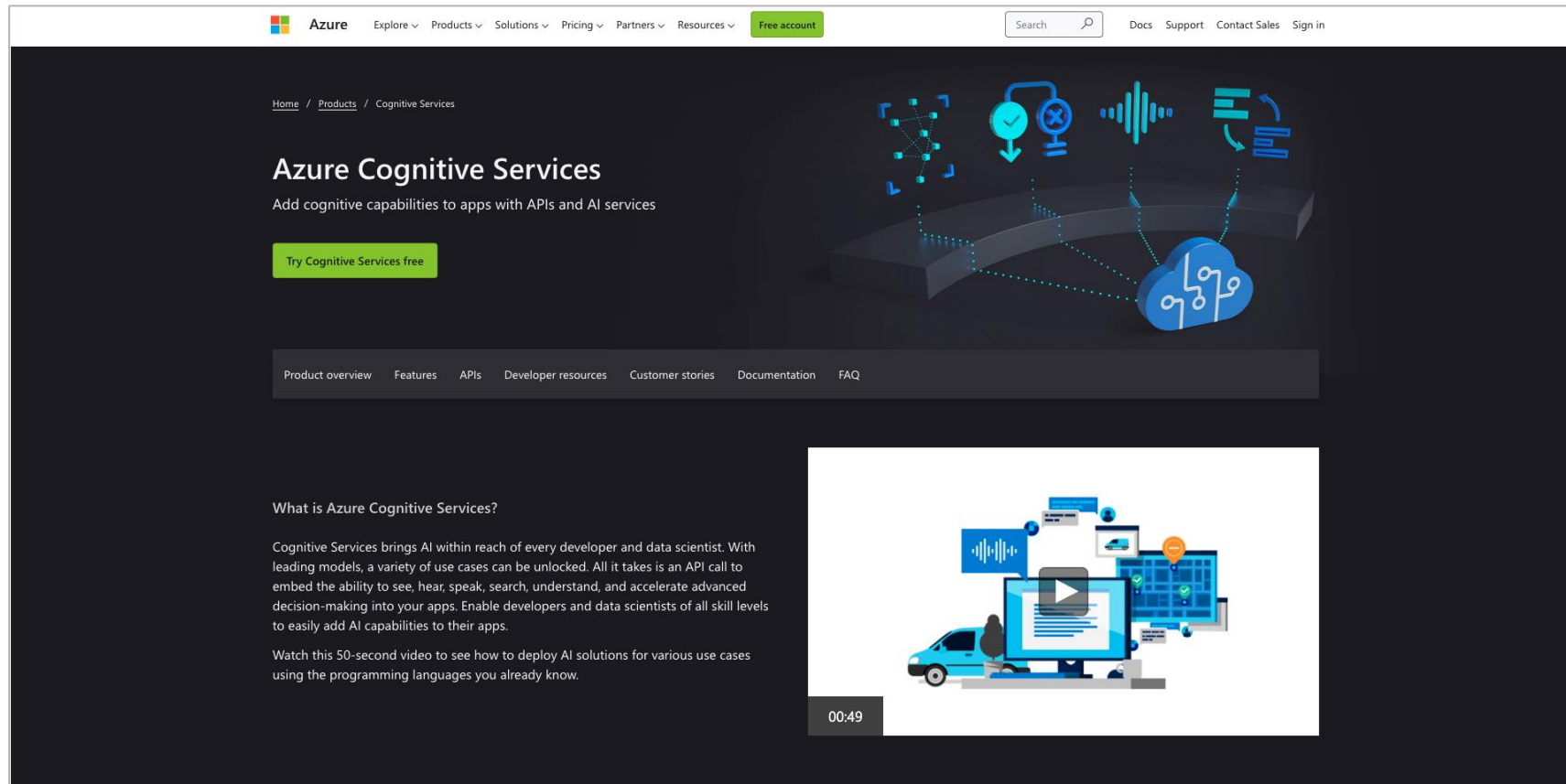
- <https://cloud.google.com/products/machine-learning/?hl=ja>
- 動画解析、画像認識、音声認識、テキスト分析などスタンダードなものは揃う
- 翻訳や画像認識、文字認識は高精度なことで評判



ディープラーニング系主要プラットフォーム（4/4）

■ Azure Cognitive Services

- <https://azure.microsoft.com/ja-jp/services/cognitive-services/>
- MicrosoftのクラウドAIプラットフォーム、他と同じく一通り揃っている
- OpenAI との連携によるAPI提供も開始



クラウドプラットフォームの比較

	音声	言語	画像	料金 (画像で比較)	備考
AWS AI	×	×	○	\$1.0/1000pic	安いが日本語非対応
Google Cloud ML	◎	○	◎	\$1.5/1000pic	翻訳、画像認識、音声認識精度に優れる
Azure Cognitive Service	○	○	◎	\$0.9/1000pic	APIが充実、検索系が強い
IBM Watson	○	◎	○	\$2.0/1000pic	高めだが日本語処理に強い

クラウドやAPIプラットフォームを使う時の注意点

- データの利用範囲・帰属などに十分注意する（利用規約を良く読む）
 - リクルートAPIのように社内利用するという旨が記載されている場合もある
 - 「使います」と明言せず分かりにくく書いているケースが多い
- 無償APIの場合、動作保証が無いことが前提
 - 意図しない動作やダウンが発生しても責任を問えない
 - リクエストが多すぎるとレスポンスが遅くなったり制限される場合も