

計算機による 学習理論

次世代IT人材育成セミナー

鹿児島大学大学院 理工学研究科
准教授 渕田孝康

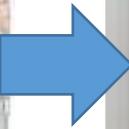
ネコを認識する計算機？



- 2012年6月26日
 - Google社の人工知能が、人間に教わることなく**自力**でネコの顔を認識することに成功した、と発表。
(<https://googleblog.blogspot.jp/2012/06/using-large-scale-brain-simulations-for.html>)
- Deep Learning(深層学習)
 - YouTubeからランダムに1000万枚の200x200画像
 - 1000台の計算機(16,000コア)で3日間かけて学習
 - ただ画像を見せるだけ
 - 画像が何か、は教えない(使わない)
 - 特定の画像に強く反応するニューロンの創生

原論文

- “Building High-level Features Using Large Scale Unsupervised Learning”, Proc. of 29th ICML, 2012
 - 人の顔、ネコの顔、人の身体を識別できた



最も反応する顔

AlphaGoがイ・セドルに勝利

- Google DeepMind社が開発した囲碁の人工知能
- 2015年10月
 - ヨーロッパ王者(プロ二段)の樊麾(Fan Hui)を5-0で破った。
 - AIがプロ棋士をハンディキャップなしの19路盤で破ったのはこれが初めて。
- 2016年3月8日～15日
 - 世界王者(プロ九段)のイ・セドルを4-1で破った。
 - 围碁界ではかなり衝撃的
- 2017年5月23日～27日
 - 世界最強棋士(プロ九段)のカ・ケツを3-0破った。
 - AlphaGO引退表明

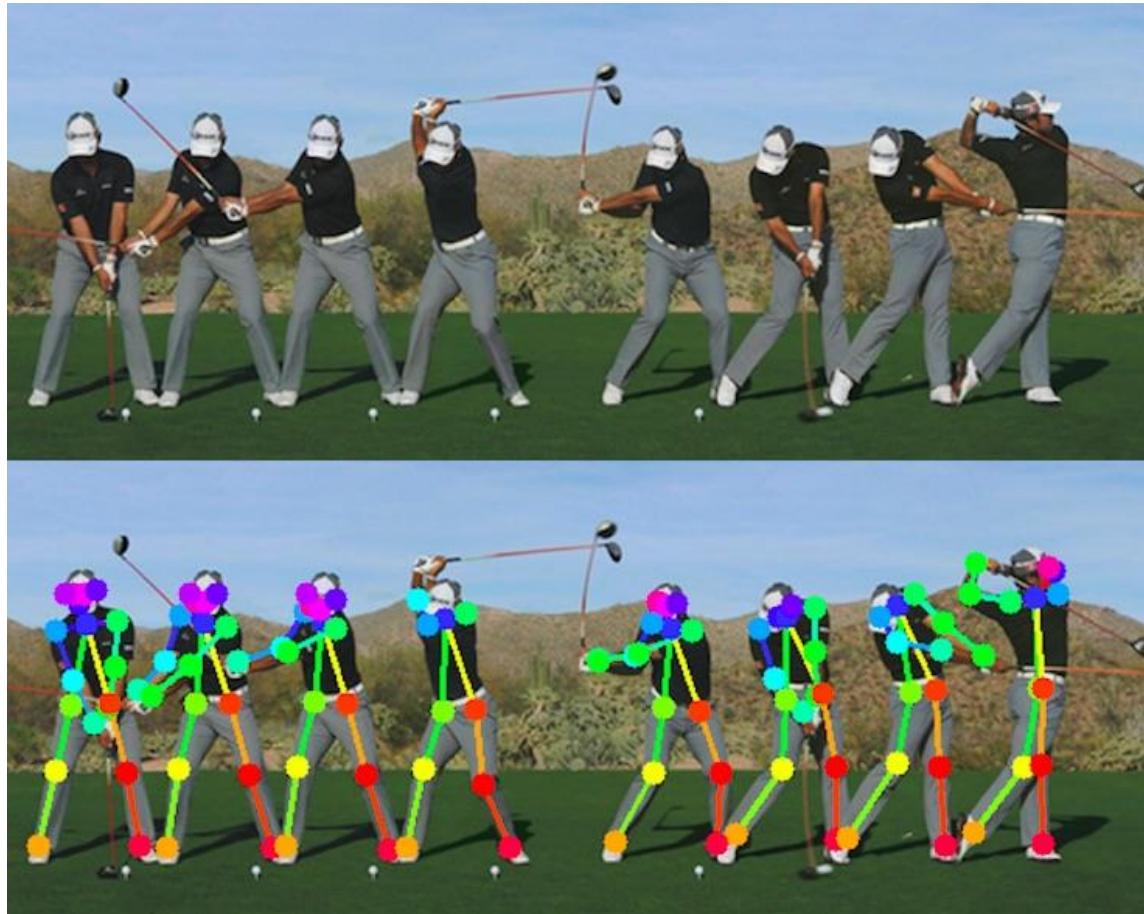


新华网
WWW.NEWS.CN

OpenPose (2017.Apr)

- カーネギーメロン大学(CMU)の Zhe Caoら が「Realtime Multi-Person pose estimation」の論文で発表
- 静止画を入力するだけで人間の関節点を検出
- 動画像内に複数人の人物がいても、リアルタイムに検出することも可能(GPU使用)
- 身体だけでなく、顔と手まで解析可能
- https://github.com/ZheC/Realtime_Multi-Person_Pose_Estimation

静止画からの骨格検出

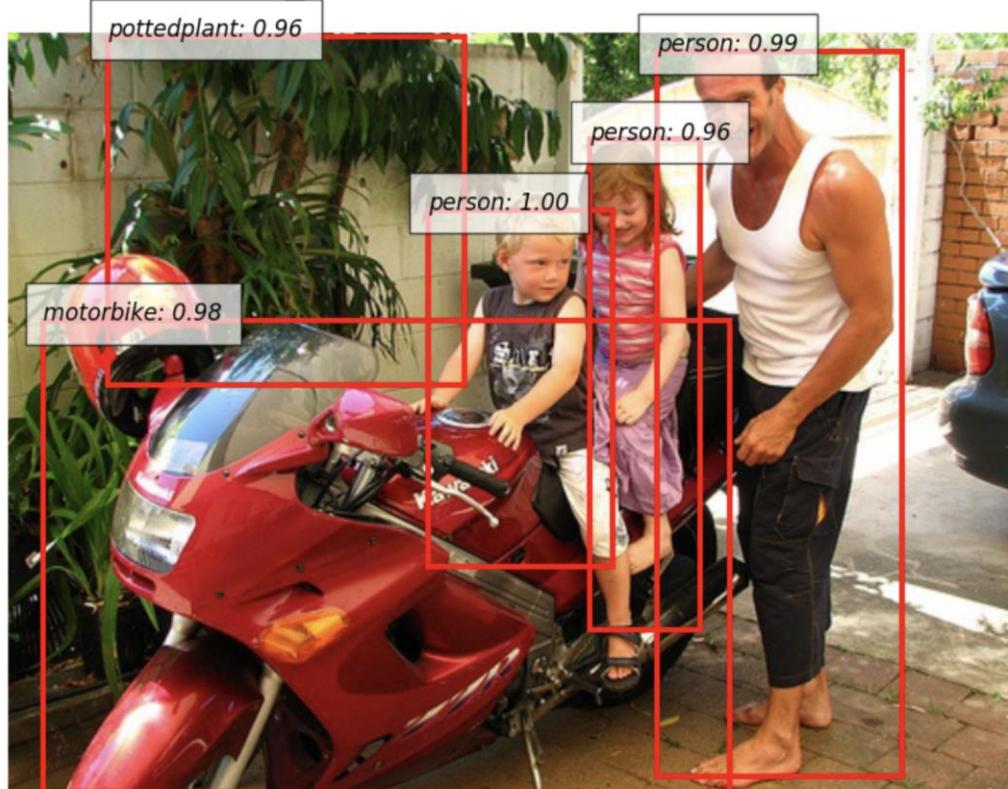


動画からの骨格検出



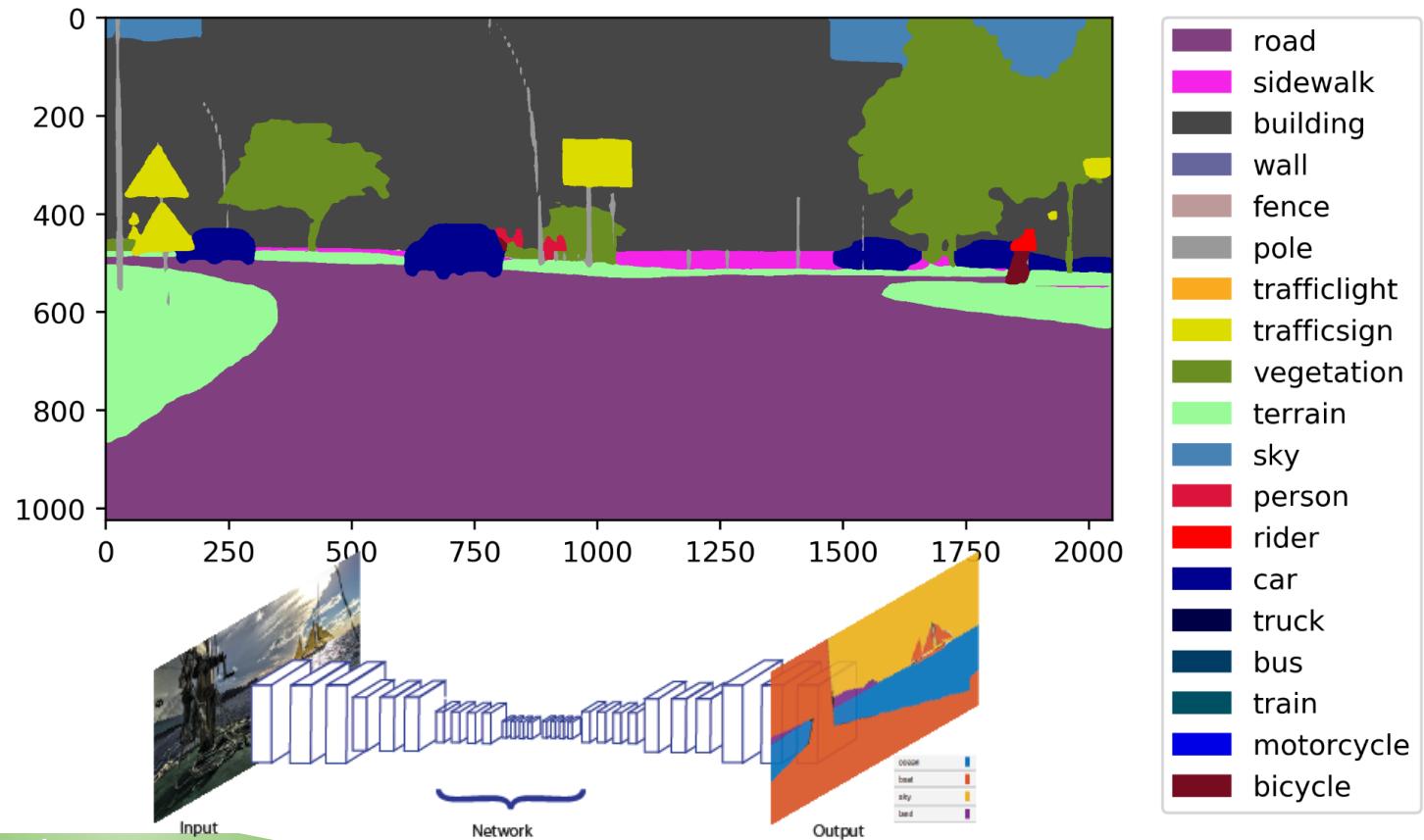
セグメンテーション

- ・物体検出
 - ・矩形枠で、どこに何があるかを識別



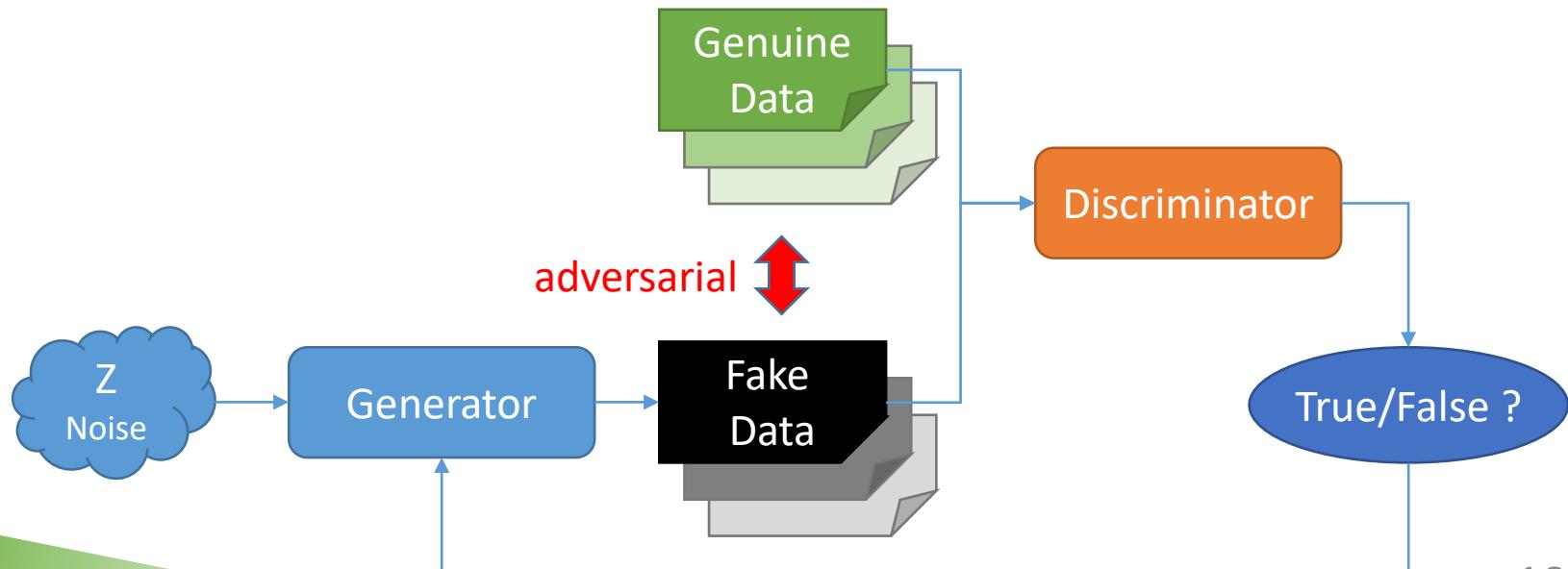
セグメンテーション

- セマンティックセグメンテーション
 - イメージをピクセルレベルとクラスラベルに関連付ける



GAN(敵対的生成ネットワーク)

- ・情報生成型ニューラルネットワーク
 - ・現実には存在しない情報を生成することが可能
 - ・深層学習に必要な膨大な学習用データを生成
 - ・存在しない人の顔を生成



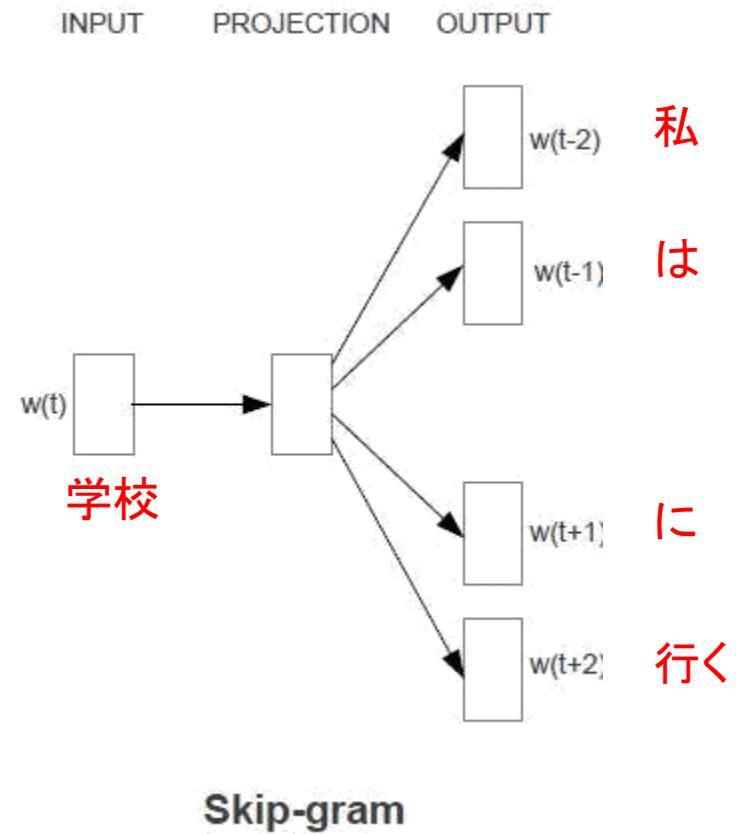
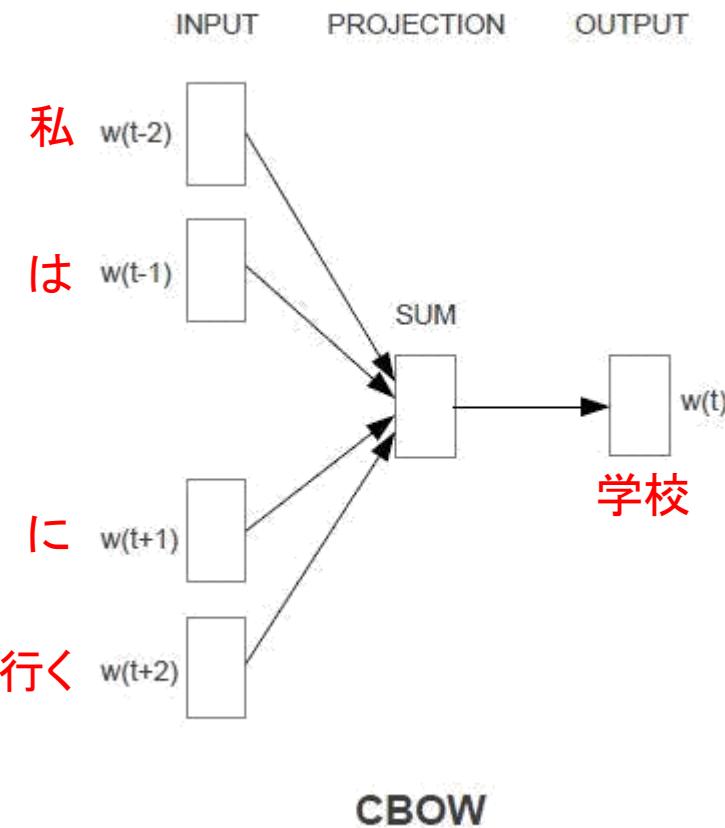
GAN情報

- GAN:敵対的生成ネットワークとは何か
- GAN(敵対的生成ネットワーク)とは
- 創造的AIと敵対的AIの不思議な関係、そしてアイデンティティへの脅威
- NVIDIA's AI Creates Beautiful Images From Your Sketches
 - Semantic Image Synthesis with Spatially-Adaptive Normalization
 - Online DEMO
 - NVIDIA playground

Word2Vec

- Tomas Mikolov, Kai Chen, Greg Corrado, Jeffrey Dean, 2013, ICLR
- 単語をベクトルに変換
- 自然言語処理における新しい手法
 - 単語で演算が可能
 - 王 - 男 + 女 = 女王
 - パリ - フランス + 日本 = 東京
- 自然言語の連なり関係から単語の類似性を抽出する

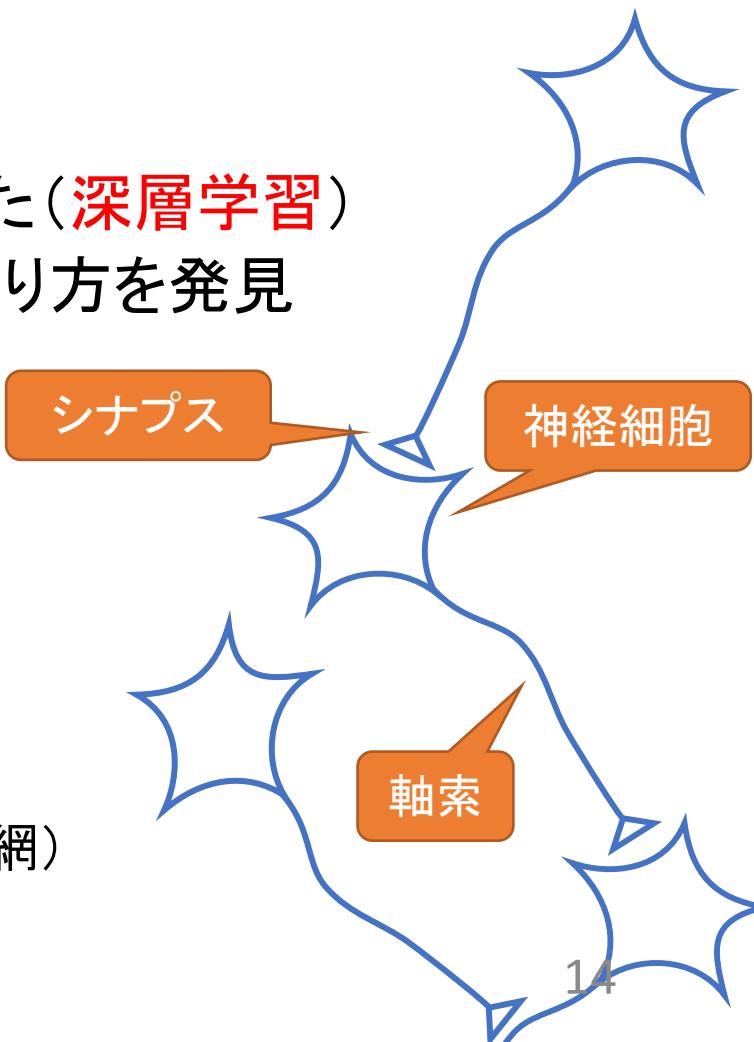
CBOWとSkip-gram



https://deepage.net/bigdata/machine_learning/2016/09/02/word2vec_power_of_word_vector.html

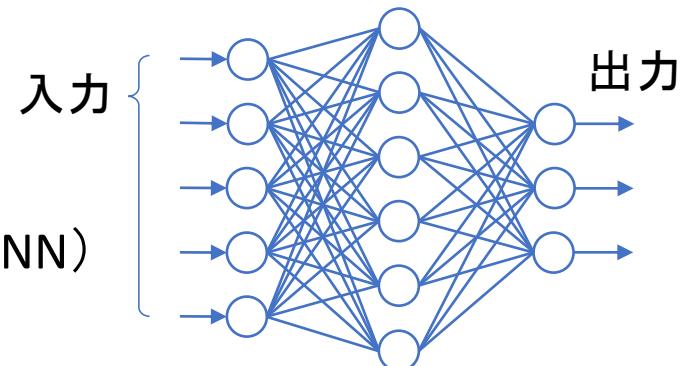
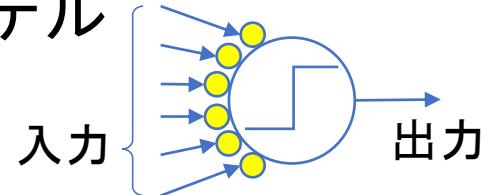
Deep Learning(深層学習)

- ネコの認識、囲碁
 - 人間はやり方を教えていない
 - 学習するための方法を作りだした(**深層学習**)
 - **膨大なサンプル**から自動的にやり方を発見
- どんな方法?
 - 人間の脳
 - 数百億個のニューロン(神経細胞)
 - 数百兆個のシナプス(神経結合)
 - その結合により情報を処理
 - 脳を真似た構造
 - ニューラルネットワーク(神経回路網)



ニューラルネットワーク

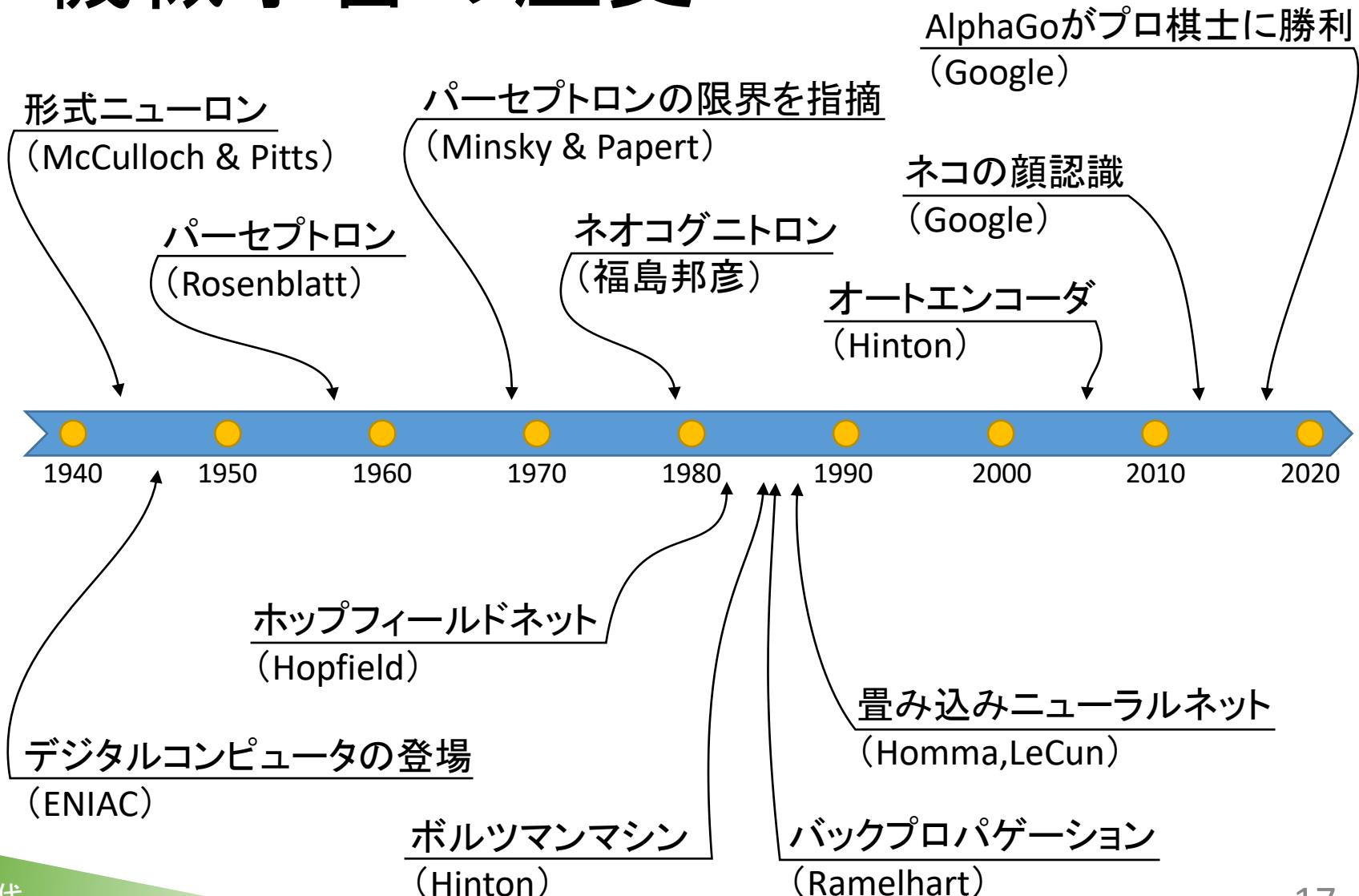
- ・形式ニューロン(McCulloch & Pitts,1943)
 - ・生物のニューロンを形式的に表現したモデル
 - ・人工ニューロンともいう
- ・ニューラルネットワーク
 - ・脳機能の特性を計算機上に表現する数理モデル
 - ・形式ニューロンのようなノードを多数個つないだもの
 - ・フィードフォワードネットワーク
 - ・パーセptron
 - ・リカレントネットワーク
 - ・ホップフィールドネットワーク
 - ・畳み込みニューラルネットワーク(CNN)



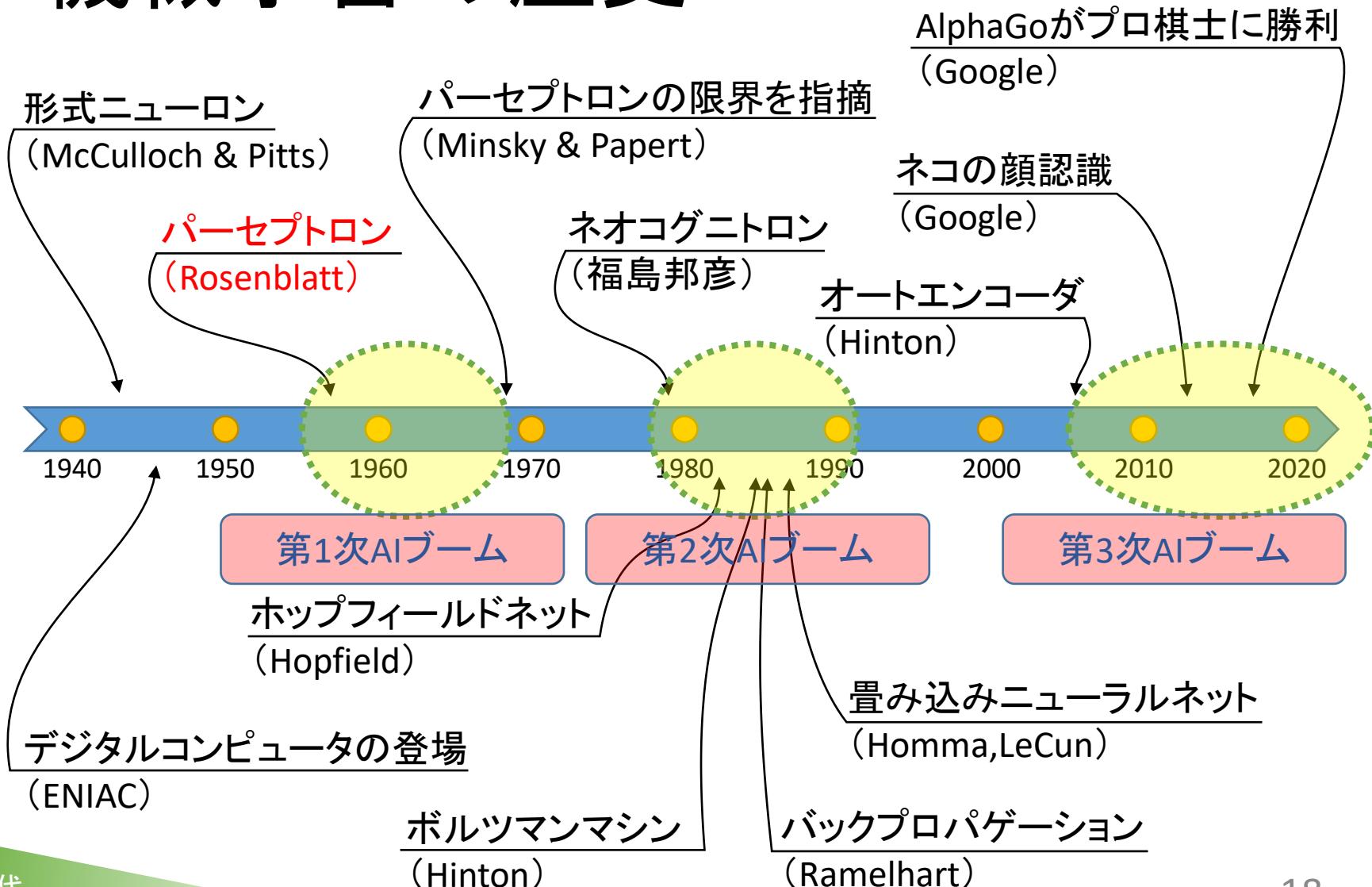
機械学習とは？

- 人間が自然に行っている学習能力と同様の機能をコンピュータで実現しようとする技術・手法（ワイキペディア）
- 機械学習の種類
 - 教師あり学習（Supervised Learning）
 - 教師なし学習（Unsupervised Learning）
 - 強化学習（Reinforcement Learning）
- 実現技法
 - ニューラルネットワーク、遺伝的アルゴリズム、相關学習、サポートベクターマシン、クラスタリング、Q学習

機械学習の歴史

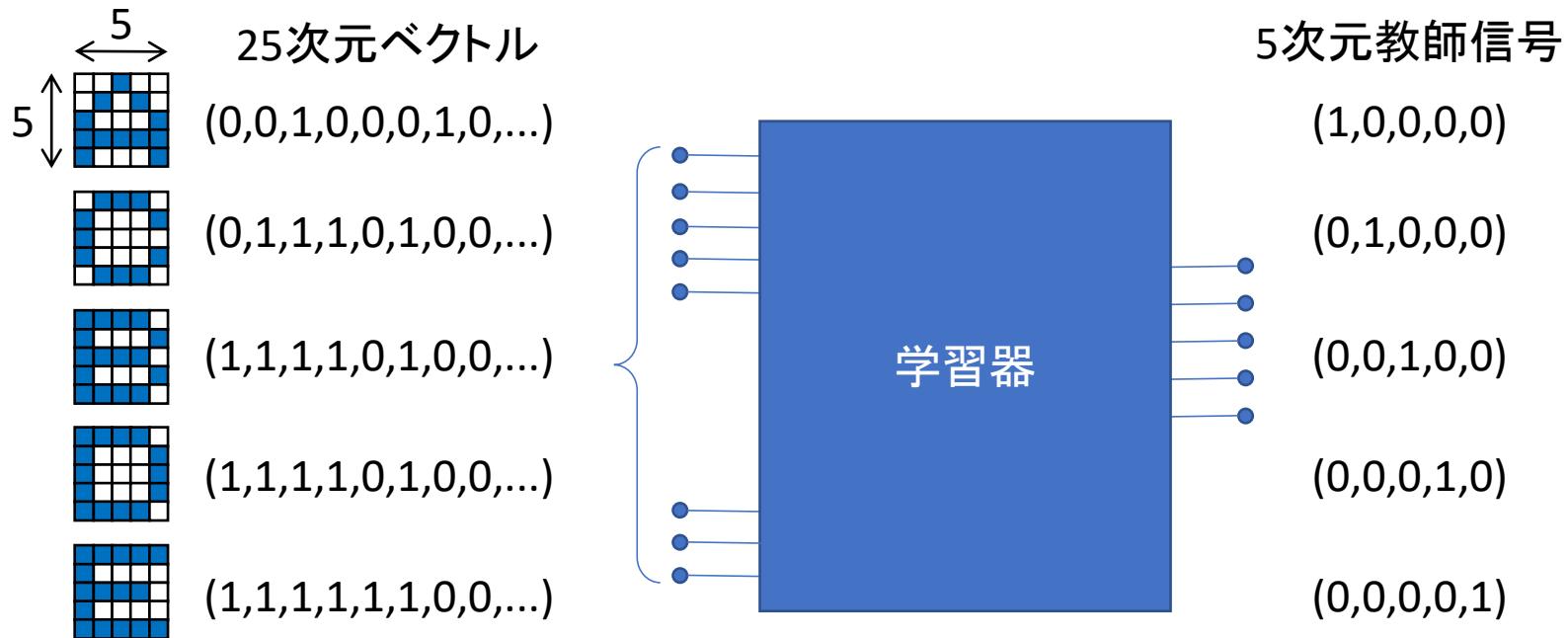


機械学習の歴史



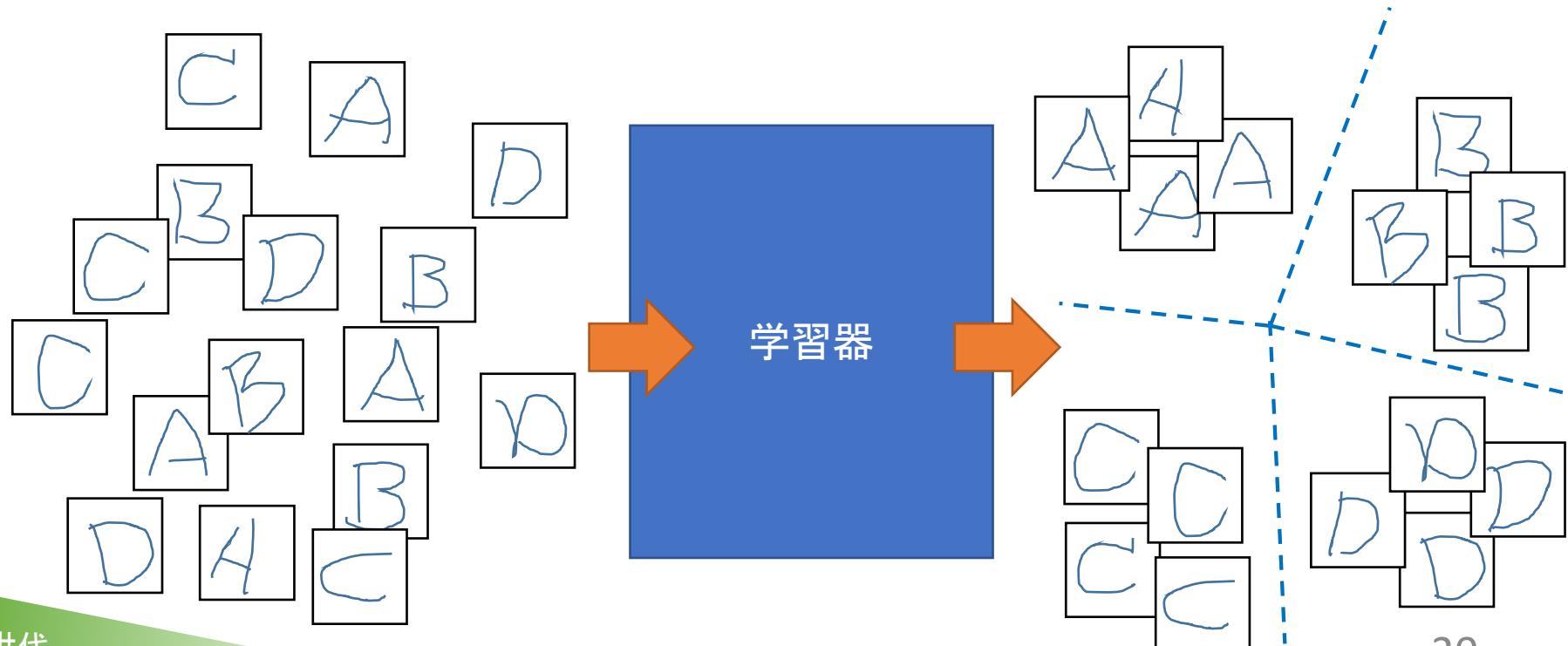
教師あり学習 (Supervised Learning)

- ・入力と教師信号の組を与える
- ・出力を教師信号に近づけるように学習する



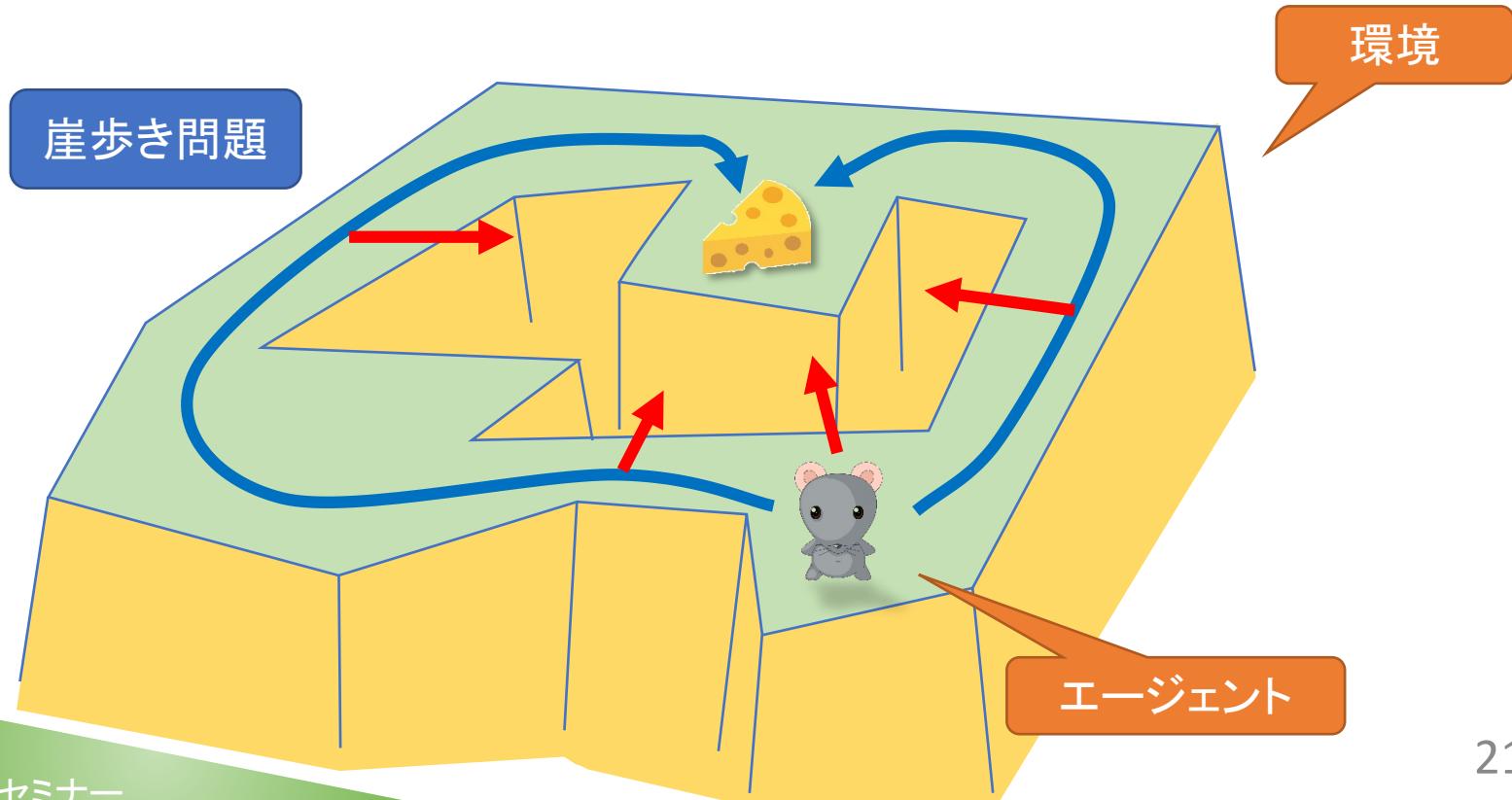
教師なし学習 (Unsupervised Learning)

- ・入力信号だけを入れて学習する
- ・入力集合が持つ潜在的な特徴・構造を抽出する



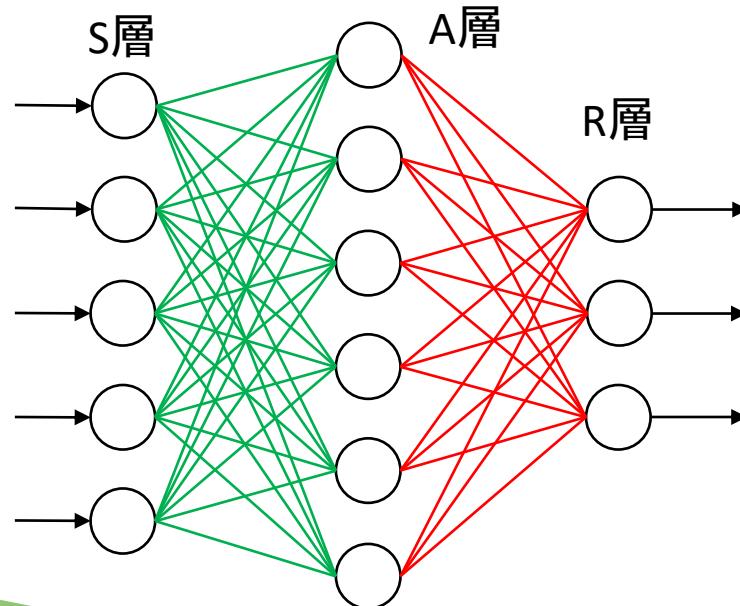
強化学習 (Reinforcement Learning)

- ・目的だけを与えて、やり方は与えない
- ・試行錯誤を繰り返してやり方を見つける



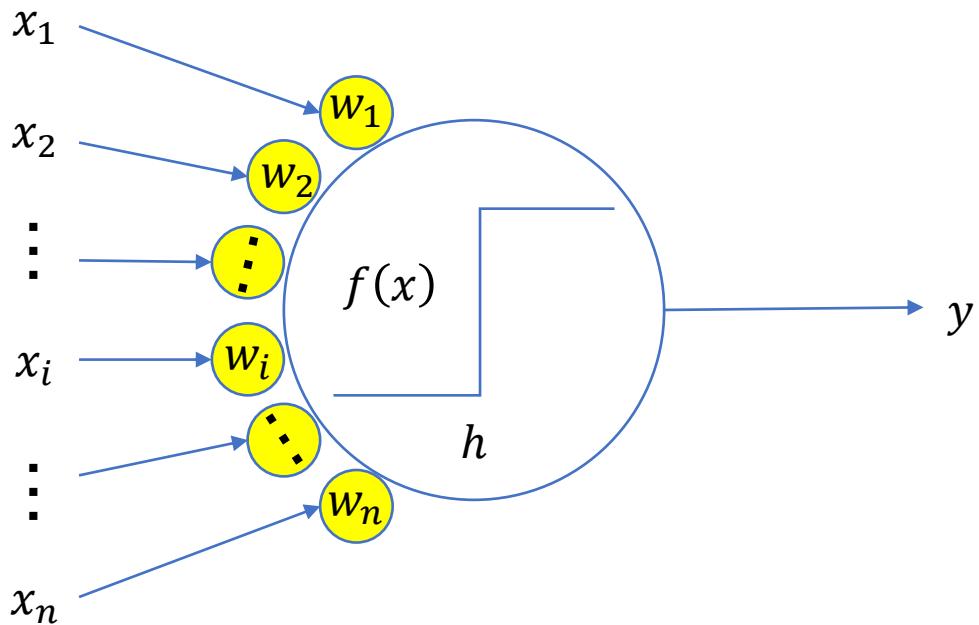
パーセプトロン

- 1957 Rosenblattらによって提案
 - S,A,Rの3層ニューラルネットワーク
 - S層(感覚・入力層)とA層(連合・中間層)の間はランダム結合
 - A層とR層(反応・出力層)間の重みを学習する



形式ニューロン

- ・入力の積和からしきい値を引いた値で出力が決まる
 - ・ $f(x)$: 階段状関数(ステップ関数)

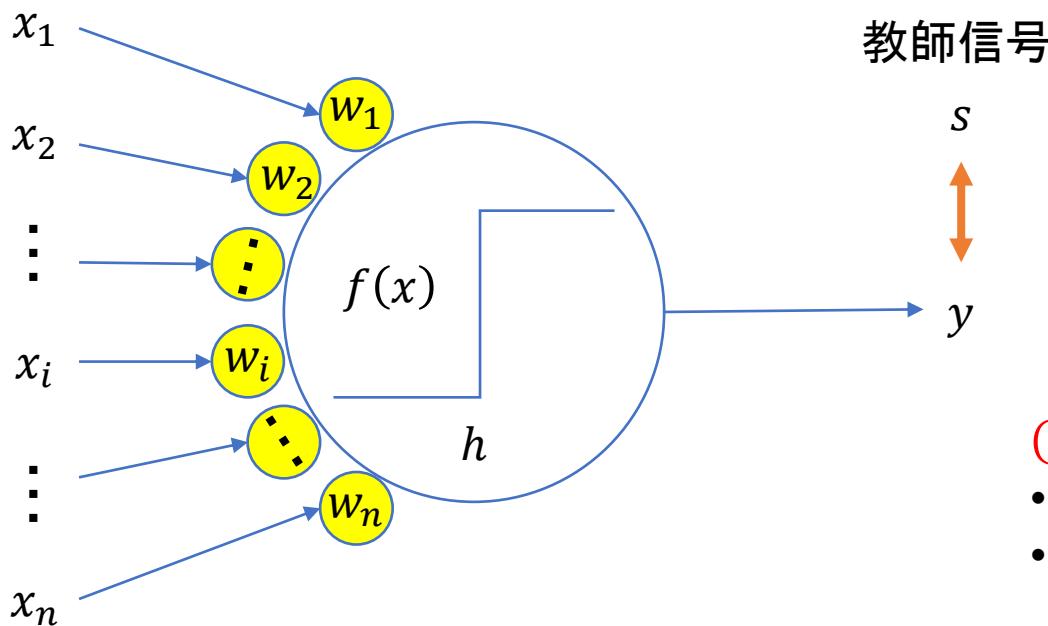


$$y = f \left(\sum_{i=1}^n w_i x_i - h \right)$$

$$f(x) = \begin{cases} 1 & (x \geq 0) \\ 0 & (x < 0) \end{cases}$$

形式ニューロンの学習

- 出力と教師信号との誤差を減らすように重みを修正する



教師信号

s

y

$$w_i = w_i + \alpha(x_i(s - y))$$

$$h = h + \alpha(s - y)$$

α : 学習率
(小さな正の数)

- $(s - y)$ が誤差 :
- 正しい出力の場合は0
 - そうでない場合は1か-1

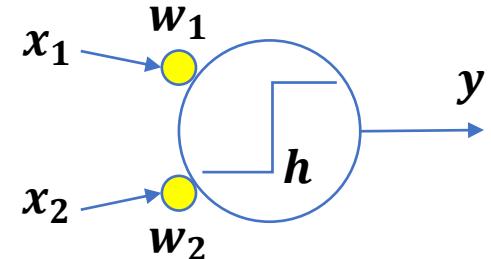
形式ニューロンの学習

- 2入力-1出力の場合

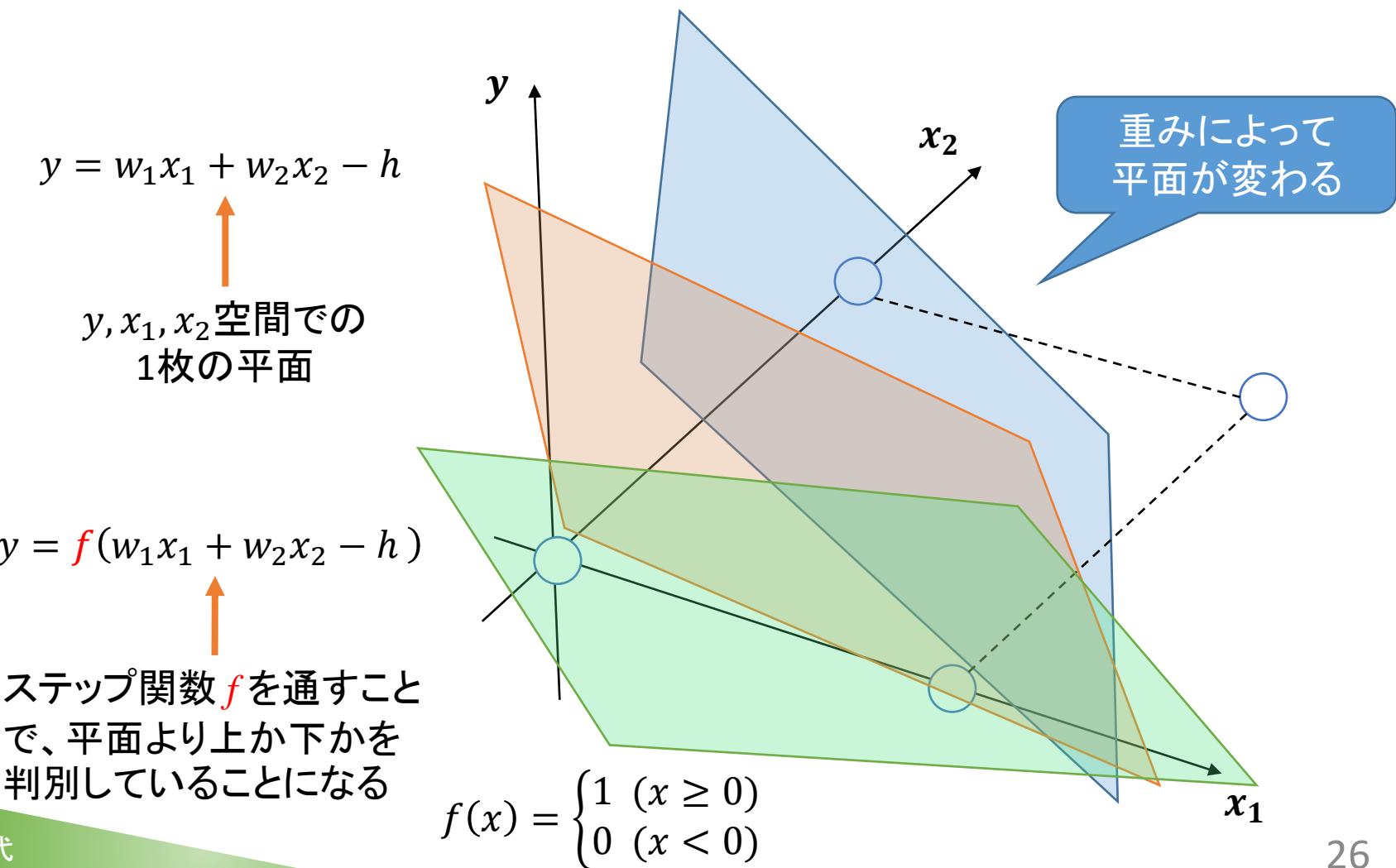
x_1	x_2	s
0	0	1
1	0	0
0	1	0
1	1	0

回数	w_1	w_2	h	x_1	x_2	$\sum_{i=1}^2 w_i x_i - h$	y	s	$s - y$
1	0	0	0	1	0	0	1	0	-1
2	-0.1	0	-0.1	0	0	-0.1	0	1	1
3	-0.1	0	0	0	1	0	1	0	-1
4	-0.1	-0.1	-0.1	1	1	-0.3	0	0	0
5	-0.1	-0.1	-0.1	0	0	-0.1	0	1	1
6	-0.1	-0.1	0	1	0	-0.1	0	0	0
7	-0.1	-0.1	0	0	0	0	1	1	0

収束



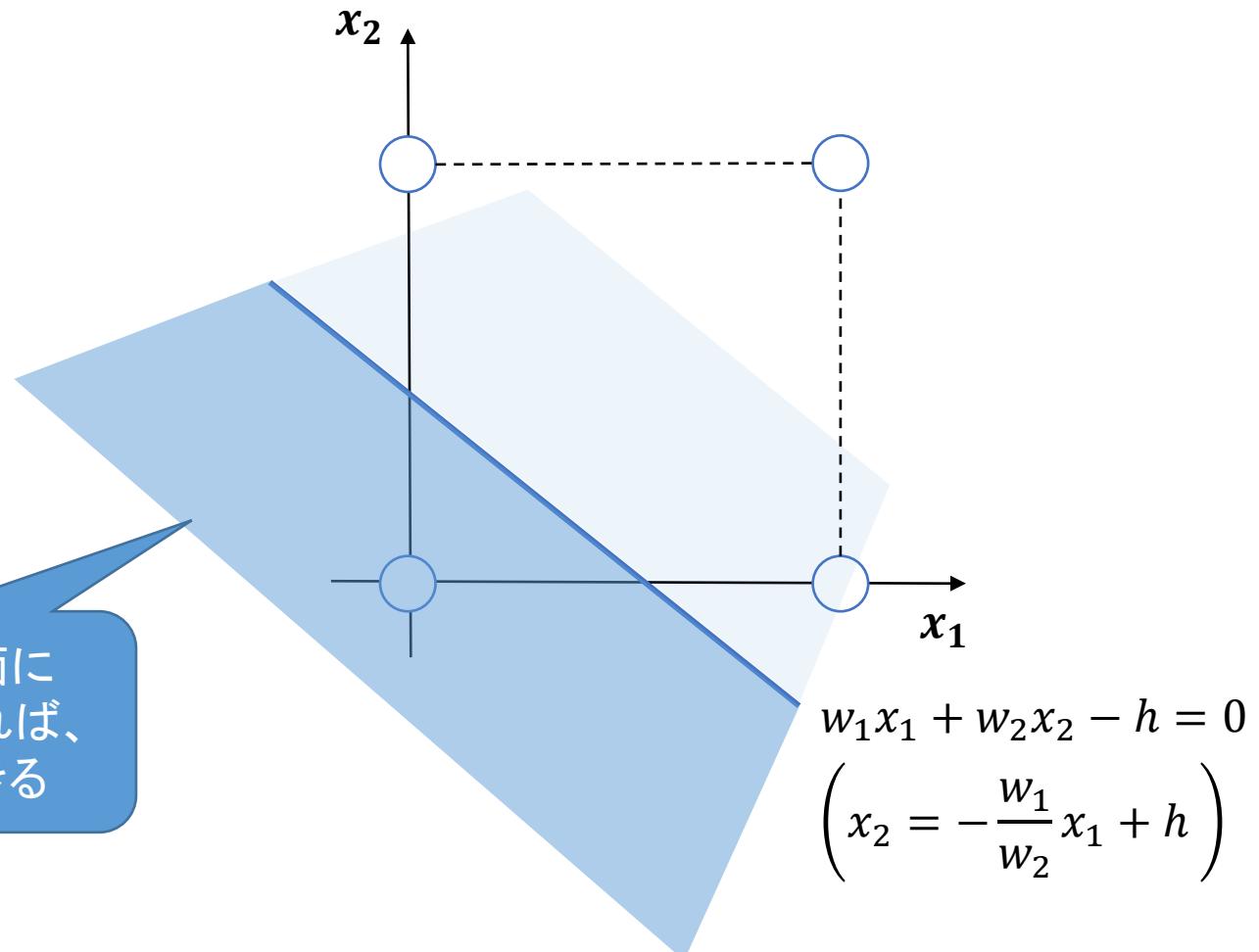
形式ニューロンの学習



形式ニューロンの学習

x_1	x_2	s
0	0	1
1	0	0
0	1	0
1	1	0

このような平面に
することができれば、
正しく判別できる

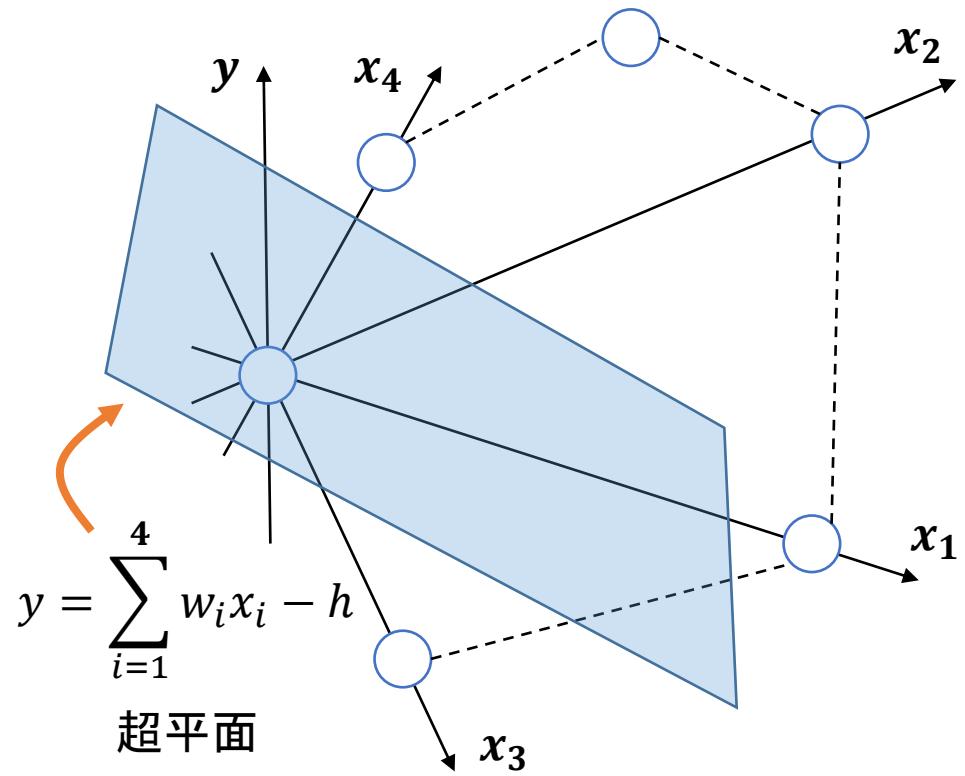


$$w_1x_1 + w_2x_2 - h = 0$$
$$\left(x_2 = -\frac{w_1}{w_2}x_1 + h \right)$$

多次元入力の場合は…

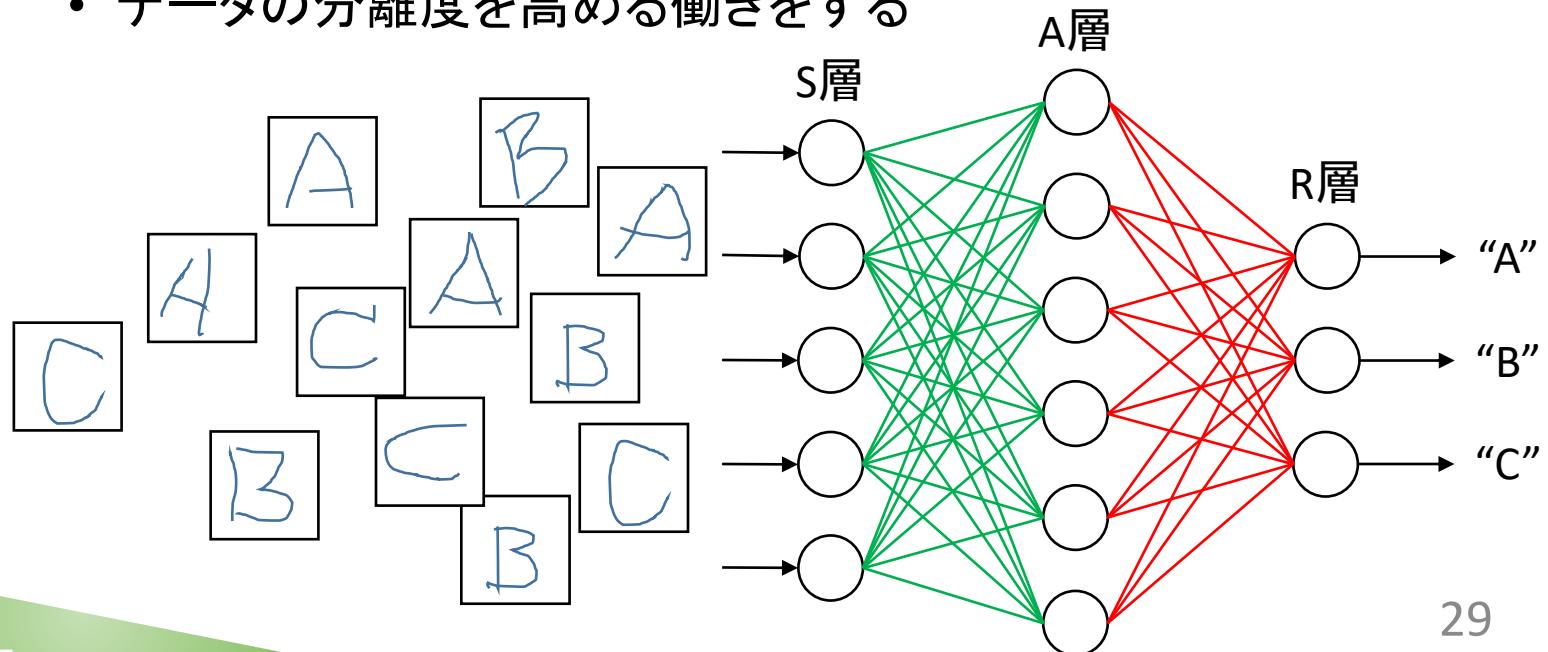
- 2入力の場合と同様だが、入力ベクトル空間が多
次元となる

x_1	x_2	x_3	x_4	s
0	0	0	0	1
1	0	0	0	0
0	1	0	0	0
1	1	0	0	1
0	0	1	0	1
:	:	:	:	:
1	1	1	0	1
1	1	1	1	0



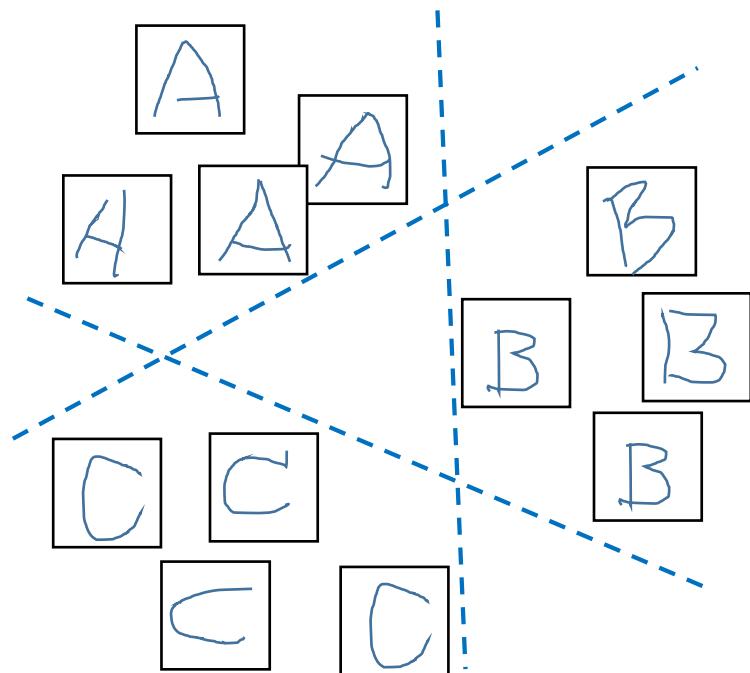
パーセプトロンの学習

- 形式ニューロンをたくさん並べたもの
 - 基本的には形式ニューロンの学習と同様
 - A層→R層の重みだけを学習する
 - S層→A層はランダム結合
 - データの分離度を高める働きをする

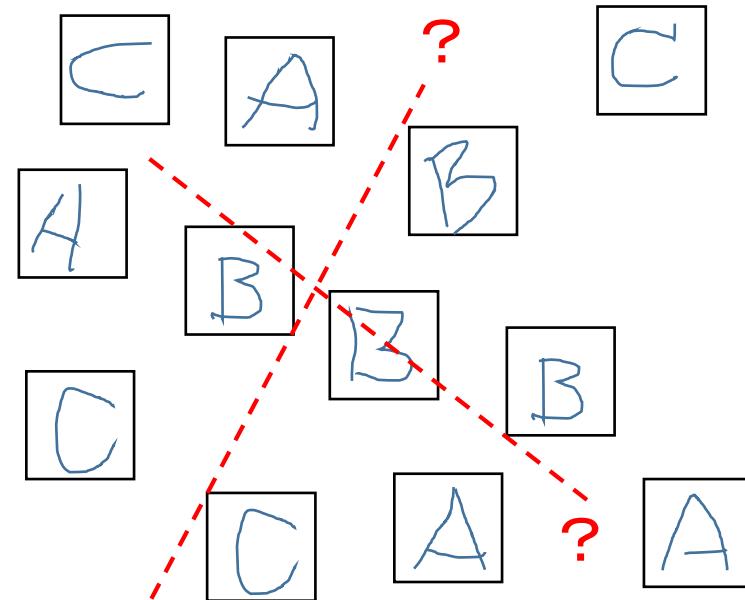


パーセプトロンの限界

- 形式ニューロンは**線形分離器**である
 - Minsky & Papert (1968)

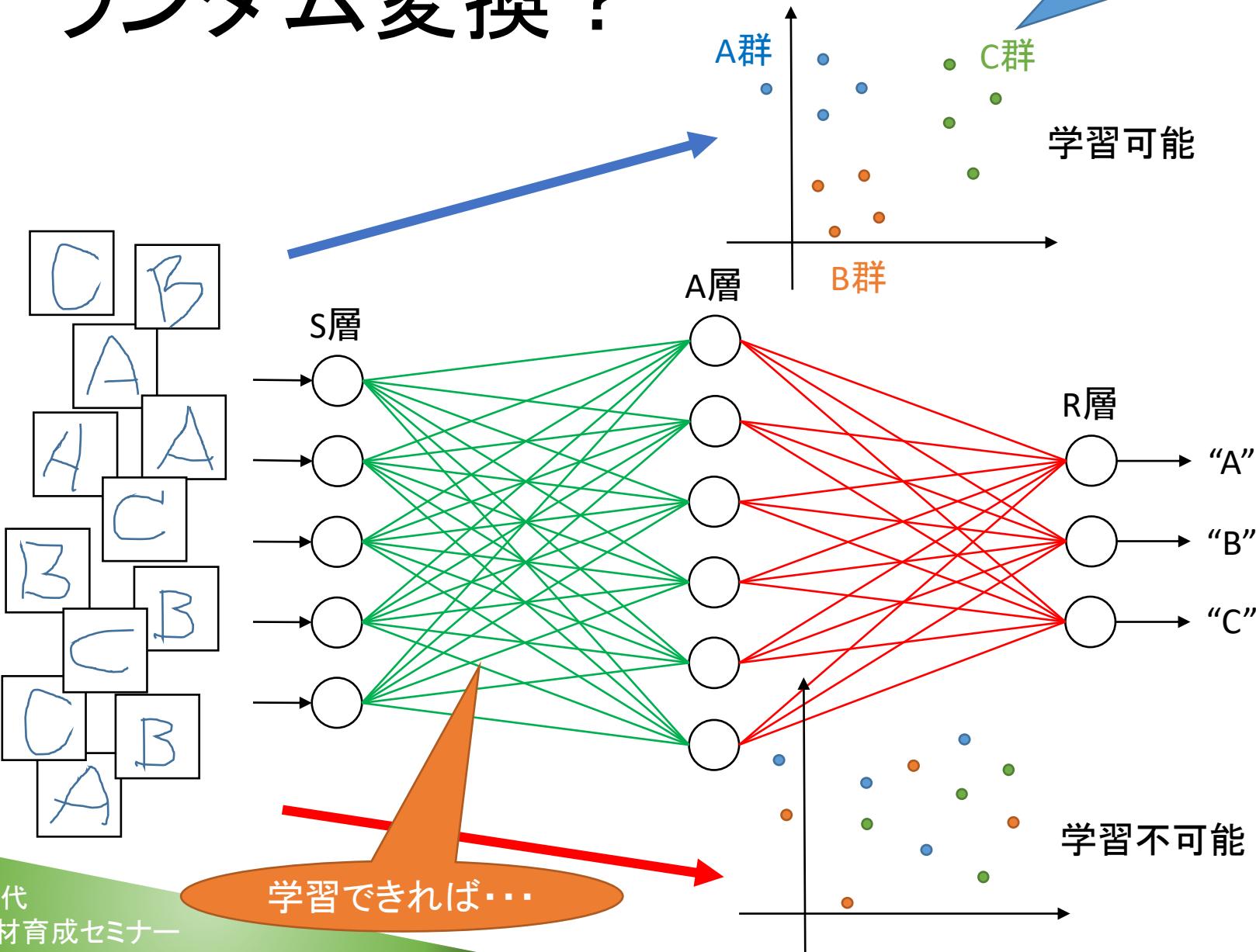


学習可能



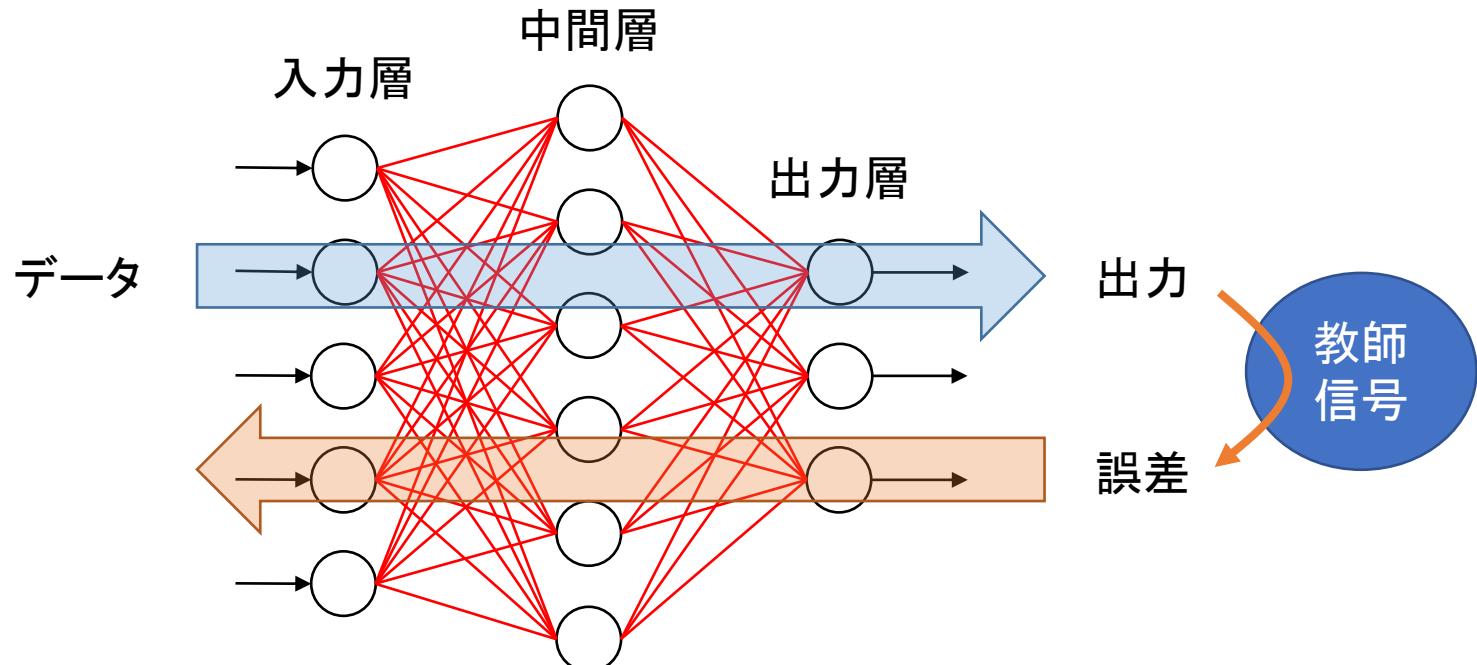
学習不可能

ランダム変換？



バックプロパゲーション学習則

- ・入力層と中間層の間の重みも学習可能
 - ・データは入力層→出力層へ伝搬
 - ・出力の誤差を出力層→入力層へ逆伝搬



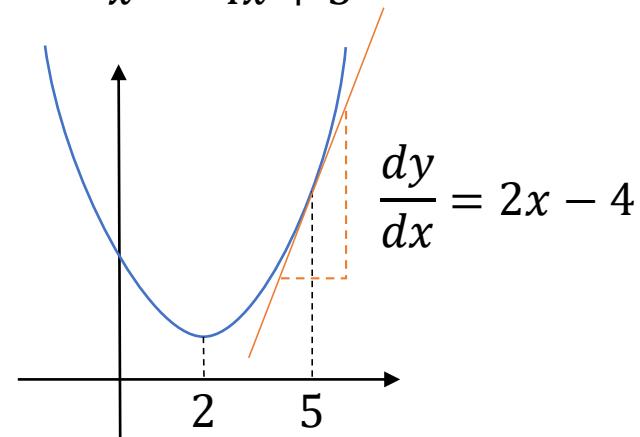
最急降下法(勾配法)

- 関数の極小値を求めたい

- 適当な数($x = 5$)から始める
- 傾きが正なので、 x を大きくすると y も大きくなる
- x から $\alpha(2x - 4)$ を引く
 - α は学習率(0.1とする)
- x は0.6減って4.4になる
- これを繰り返すと、 x は2に近づいていく

これを、ネットワークの誤差の減少に使う

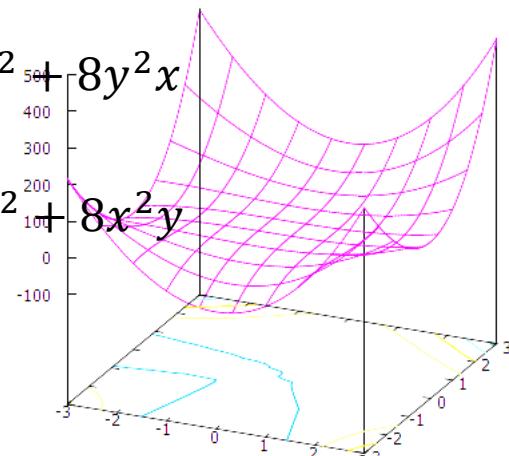
$$y = (x - 2)^2 + 1 \\ = x^2 - 4x + 5$$



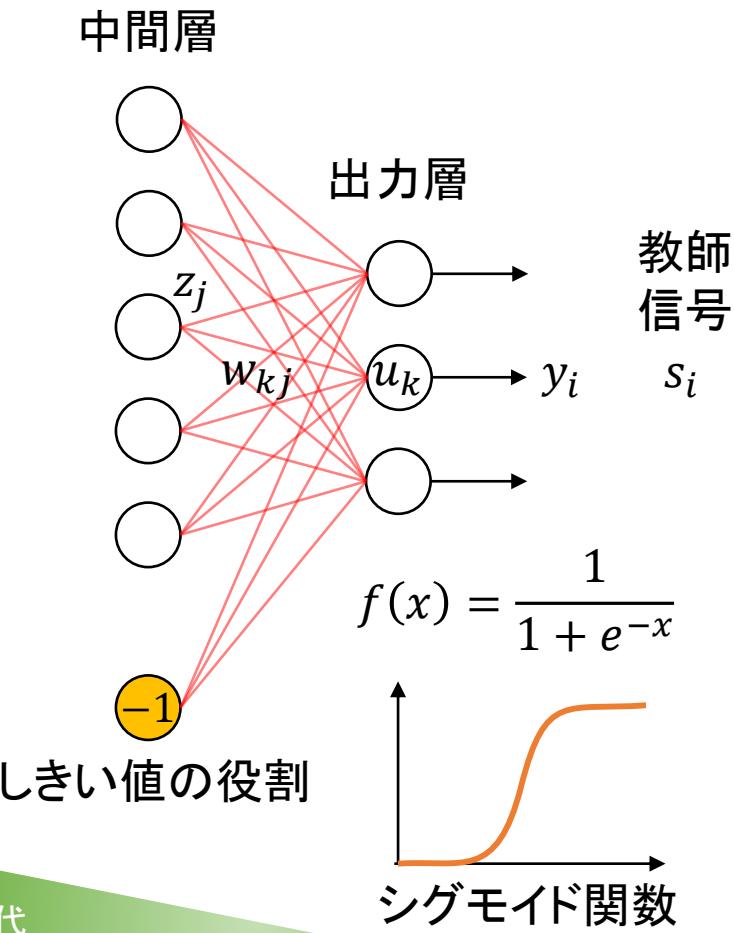
$$z = 3x^3 + 4x^2y^2 + y^3$$

$$\frac{\partial z}{\partial x} = 9x^2 + 8y^2x$$

$$\frac{\partial z}{\partial y} = 3y^2 + 8x^2y$$



出力層の学習



$$E = \frac{1}{2} \sum (s_i - y_i)^2$$

連鎖律

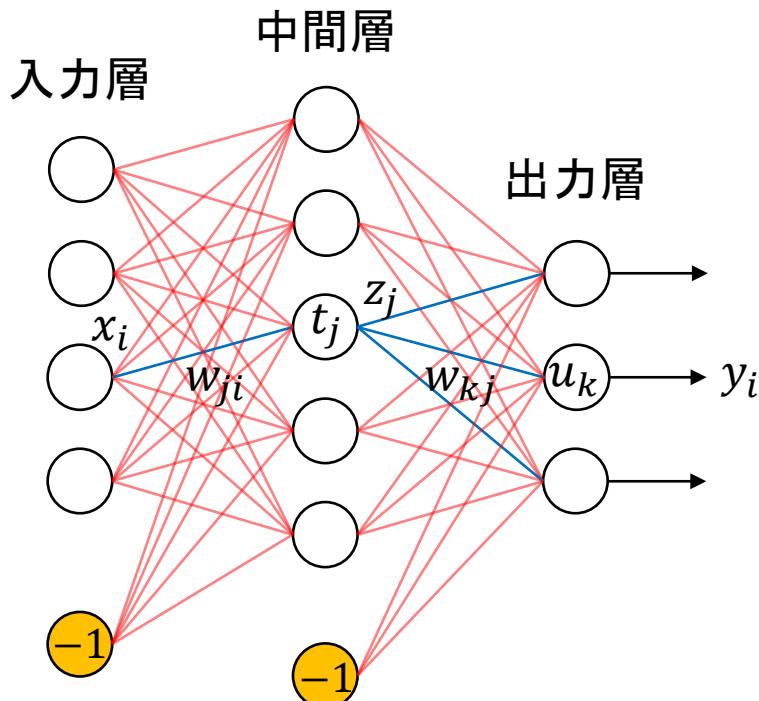
$$\Delta w_{kj} = -\alpha \frac{\partial E}{\partial w_{kj}} = -\alpha \frac{\partial E}{\partial y_k} \frac{\partial y_k}{\partial u_k} \frac{\partial u_k}{\partial w_{kj}}$$

$$\frac{\partial E}{\partial y_k} = \frac{\partial}{\partial y_k} \left(\frac{1}{2} \sum (s_i - y_i)^2 \right) = -(s_i - y_i)$$

$$\frac{\partial y_k}{\partial u_k} = \frac{\partial}{\partial u_k} \left(\frac{1}{1 + e^{-u_k}} \right) = y_k(1 - y_k)$$

$$\frac{\partial u_k}{\partial w_{kj}} = \frac{\partial}{\partial w_{kj}} \left(\sum z_j w_{kj} \right) = z_j$$

中間層の学習



$$\Delta w_{ji} = -\alpha \frac{\partial E}{\partial w_{ji}}$$

$$= -\alpha \left\{ \sum \frac{\partial E}{\partial y_k} \frac{\partial y_k}{\partial u_k} \frac{\partial u_k}{\partial z_j} \right\} \frac{\partial z_j}{\partial t_j} \frac{\partial t_j}{\partial w_{ji}}$$

$$\frac{\partial u_k}{\partial z_j} = \frac{\partial}{\partial z_j} \left(\sum z_j w_{kj} \right) = w_{kj}$$

$$\frac{\partial z_j}{\partial t_j} = \frac{\partial}{\partial t_j} \left(\frac{1}{1 + e^{-t_j}} \right) = z_j(1 - z_j)$$

$$\frac{\partial t_j}{\partial w_{ji}} = \frac{\partial}{\partial w_{ji}} \left(\sum x_i w_{ji} \right) = x_i$$

$$z = f(y_1, y_2, \dots, y_m)$$

$$y_i = g_i(x_1, x_2, \dots, x_n) \Rightarrow \frac{\partial z}{\partial x_i} = \sum_{j=1}^m \frac{\partial f}{\partial y_j} \frac{\partial y_j}{\partial x_i}$$

一般化連鎖律

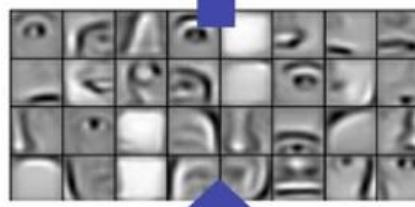
Deep Learningへ

- 4層以上の多層ニューラルネットワーク
 - 6層から20層程度
- Googleの顔認識
 - 右図のような構造の層を3つ並べて構築してある

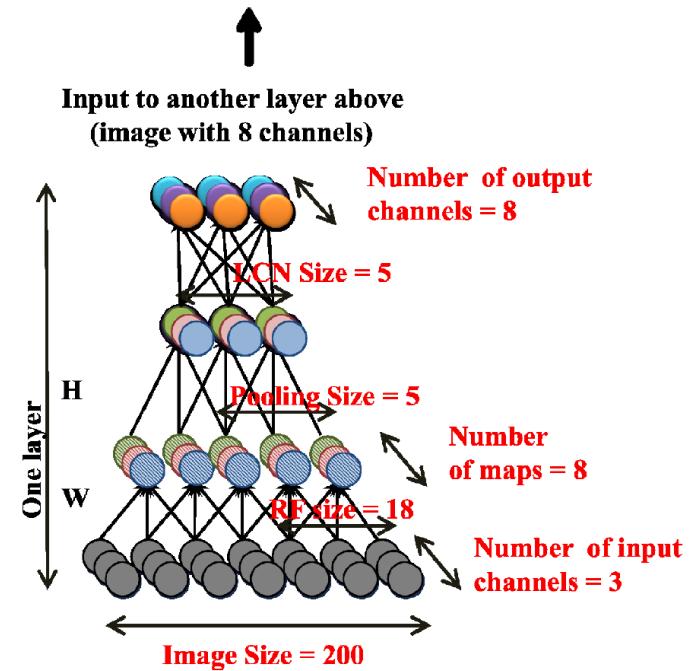
第3層



第2層



第1層



教師なし学習

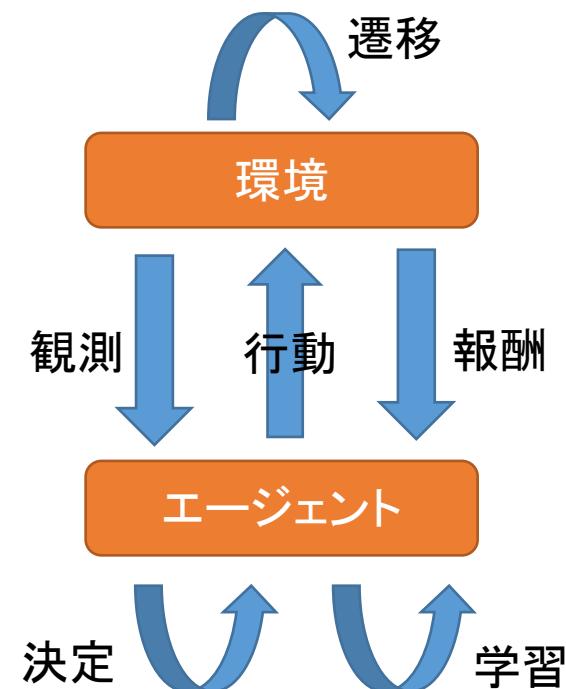
- ・正解を与えないで学習させる方法
- ・入力が持つ特徴・構造などを抽出する
- ・具体例：
 - ・クラスタリング
 - ・ベクトル量子化
 - ・自己組織化特徴マップ

K-means法

- ・クラスタリングの代表的アルゴリズム
- ・たくさんのデータをいくつかのクラスに分ける
 1. 最初にランダムにクラスを割り振る
 2. そのクラスを最も近い点とするデータを集める
 3. そのデータの平均位置へクラスを移動させる
 4. 2,3を、クラスが動かなくなるまで繰り返す
- ・デモのページ

強化学習

- ・目的だけを与えて、やり方は与えない
- ・試行錯誤を繰り返してやり方を見つける
- ・環境とエージェントとの対話
 - ・環境を観測
 - ・行動を決定
 - ・行動
 - ・環境の変化
 - ・報酬を得る
 - ・学習

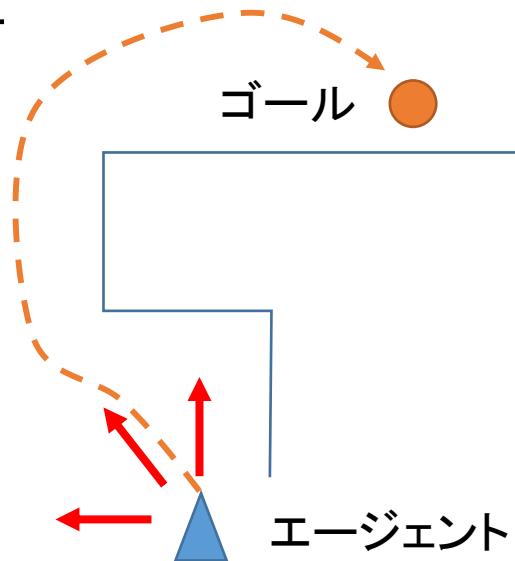


Q学習

- ・強化学習の代表的な方法
- ・報酬の期待値であるQ値を学習する
 - ・Q値:ある状態でのその行動の価値
 - ・正しいQ値が得られれば、大きなQ値の行動を選択すると、大きな報酬につながる
 - ・ランダムなQ値から始めて、正しいQ値に近づけるのが学習

$$Q_{s,a} \leftarrow Q_{s,a} + \alpha(r + \gamma \max_j Q_{s,j} - Q_{s,a})$$

割り引き率
報酬
学習率



まとめ

- 第3次AIブーム
 - 注目される深層学習
 - その原理はバックプロパゲーション
 - 教師あり学習の代表
 - 人間の脳の神経回路を真似たニューラルネットワーク
- 3つの学習理論
 - 教師あり学習
 - 教師なし学習
 - 強化学習 → 人間の脳は強化学習を行っている?
 - 楽しいことをしたい
 - 苦しいことはしたくない

これからの
研究課題