

情報学習理論



渡辺澄夫

プロローグ - 人工知能とは

実世界において自己を感じ世界を認識し
人間のように実世界を生きる<コンピュータ>
人間と同じ思考・直観・感覚・感情の機能を持つもの



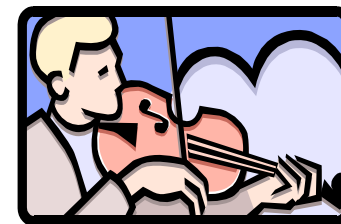
認識



予測

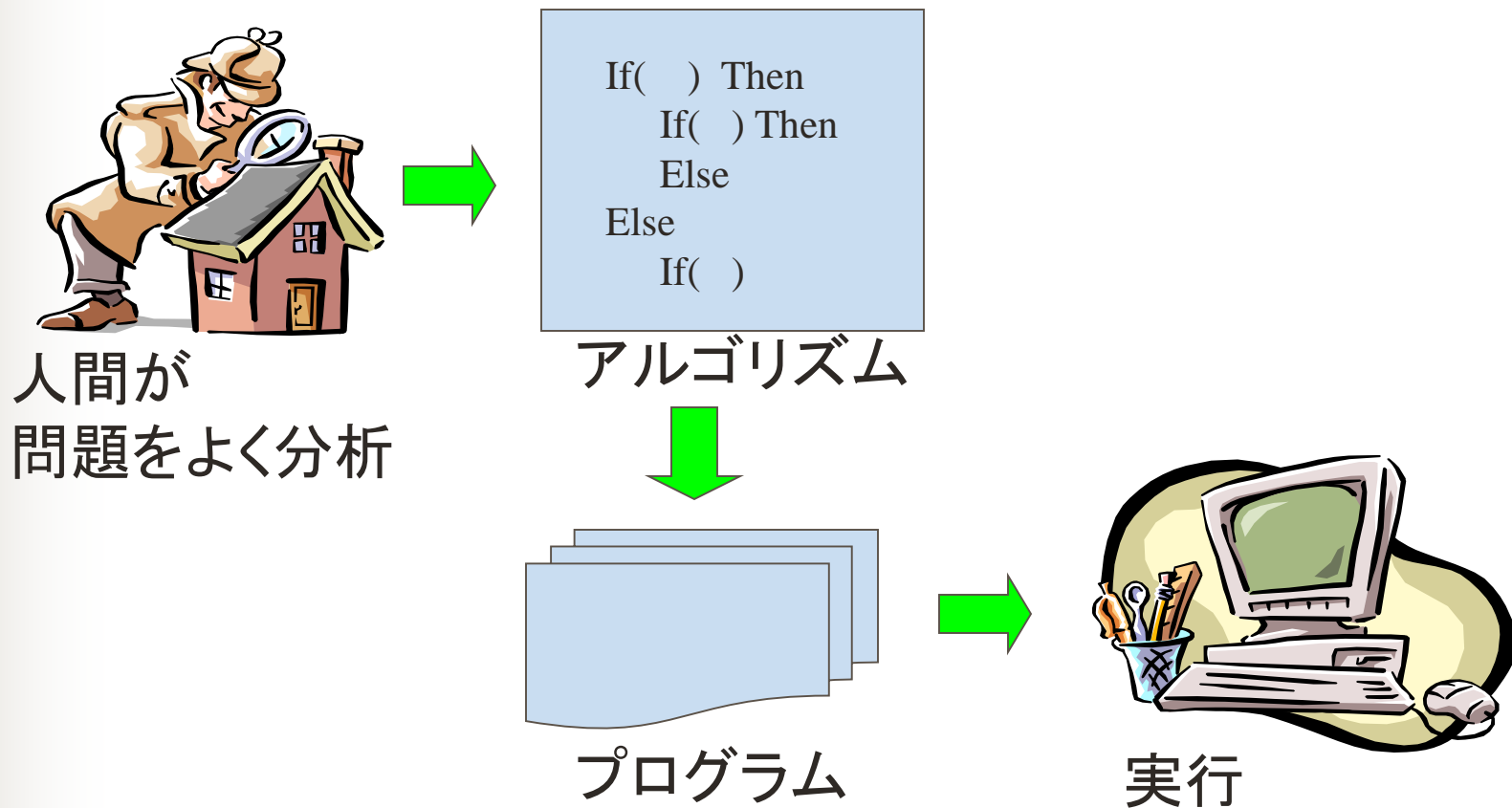


制御



創造

プロローグ - 人工知能の昔





実世界の難しさ

If Then ルールは、コンピュータの中では
うまく行ったが、実世界ではうまくいかなかった

実世界は

- 「言語化できないもの」の方が多い
- 「プログラマが予期できないこと」が起こる
- 実世界は超高次元

自ら実世界に存在することが必要

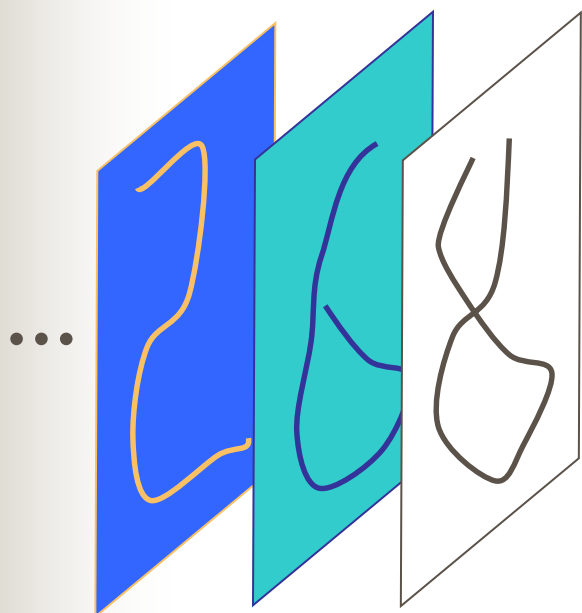
ソフトウェアの大規模化によって実証されつつある

実世界は多様

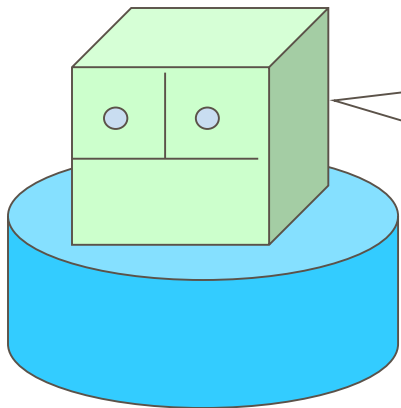


自分の力で世界を認識する

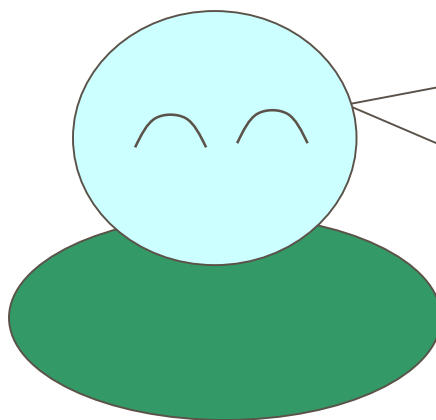
教師あり学習



文字の例



先生
8, 6, 2...
でしょう



学習者
先生と同じ
ように文字を
読みたい



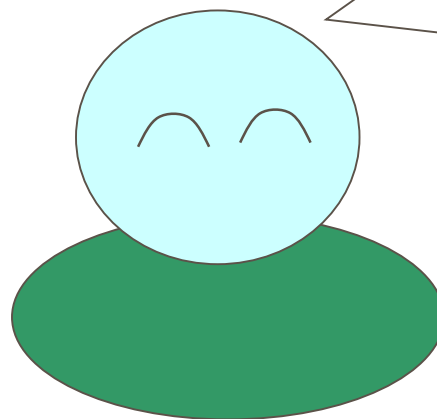
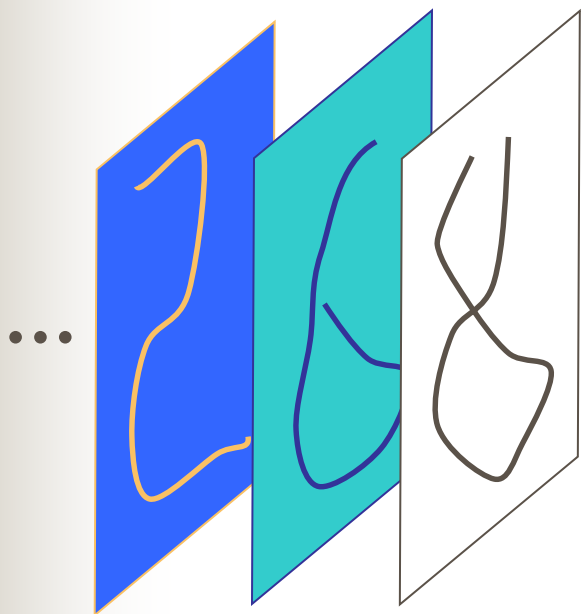
問題(1)

郵便番号として書かれた文字のデータが
100000個ある。
このデータから、文字を読みわけけるシステムを
つくるためにはどうしたらよいか？

医用画像から、自動診断あるいは診断支援を
行うためのシステムを作るにはどうしたらよいか？

ある映画に対するユーザーの評判を
Twitter のデータから推測したい。

教師なし学習



学習者

概念とは？

概念1

概念2

概念3

世界には例があふれている。例から世界の構造を知る。



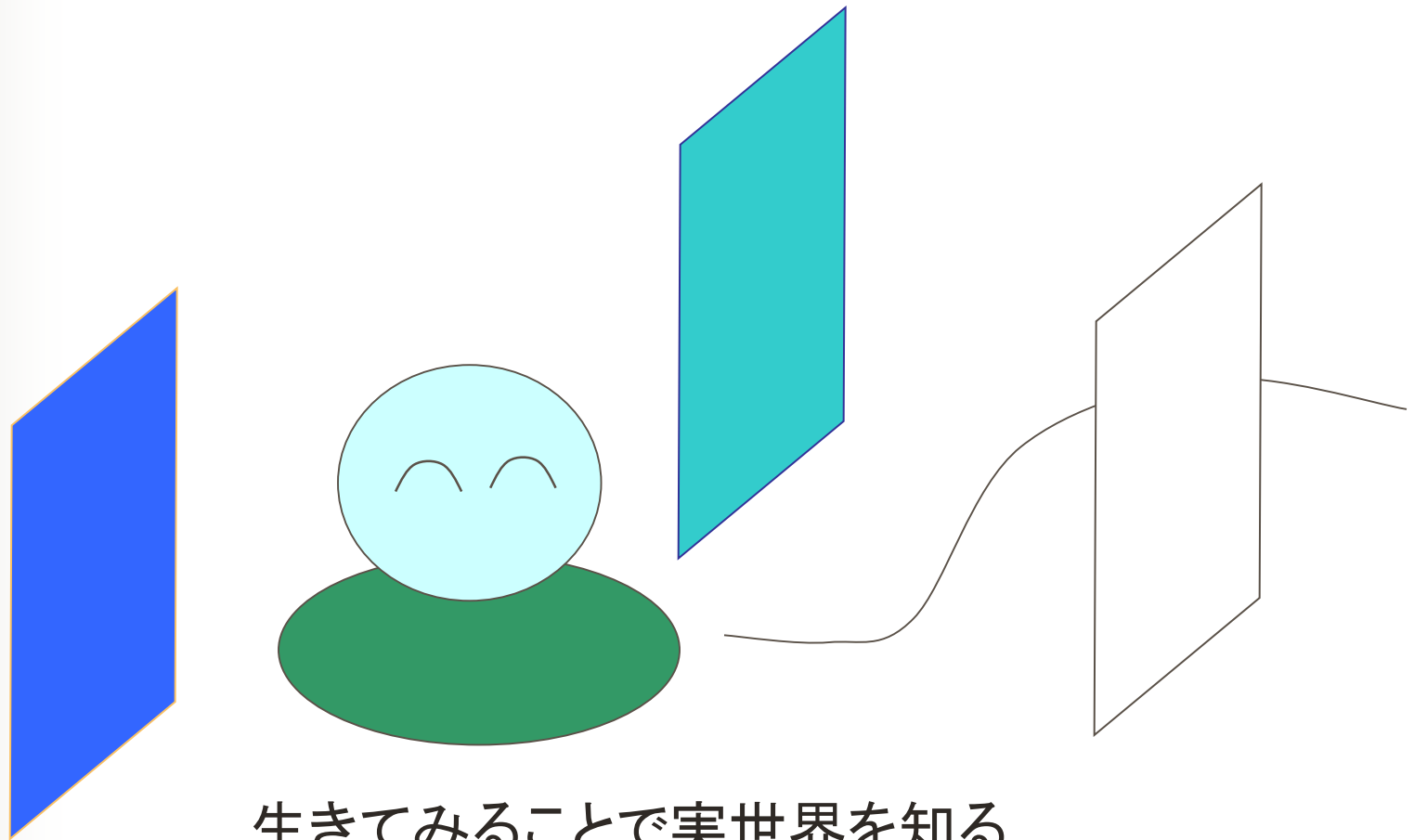
問題(2)

世界中にあるWWWページを自動的に分類し、
様々な観点から階層構造を自動生成する方法を考えよ。

自動車10000種類について、車高、車幅、長さ、馬力、
値段、・・・などのデータを集めた。これより、自動車の作る
空間の地図を作りたい。どうしたらよいか？

流行歌のデータがたくさんあるとき、そのパターンを
明らかにして、自分で作った歌が過去のどのパターンに
似ているかを判定したい。

強化学習（この講義ではやりません）



生きてみることで実世界を知る

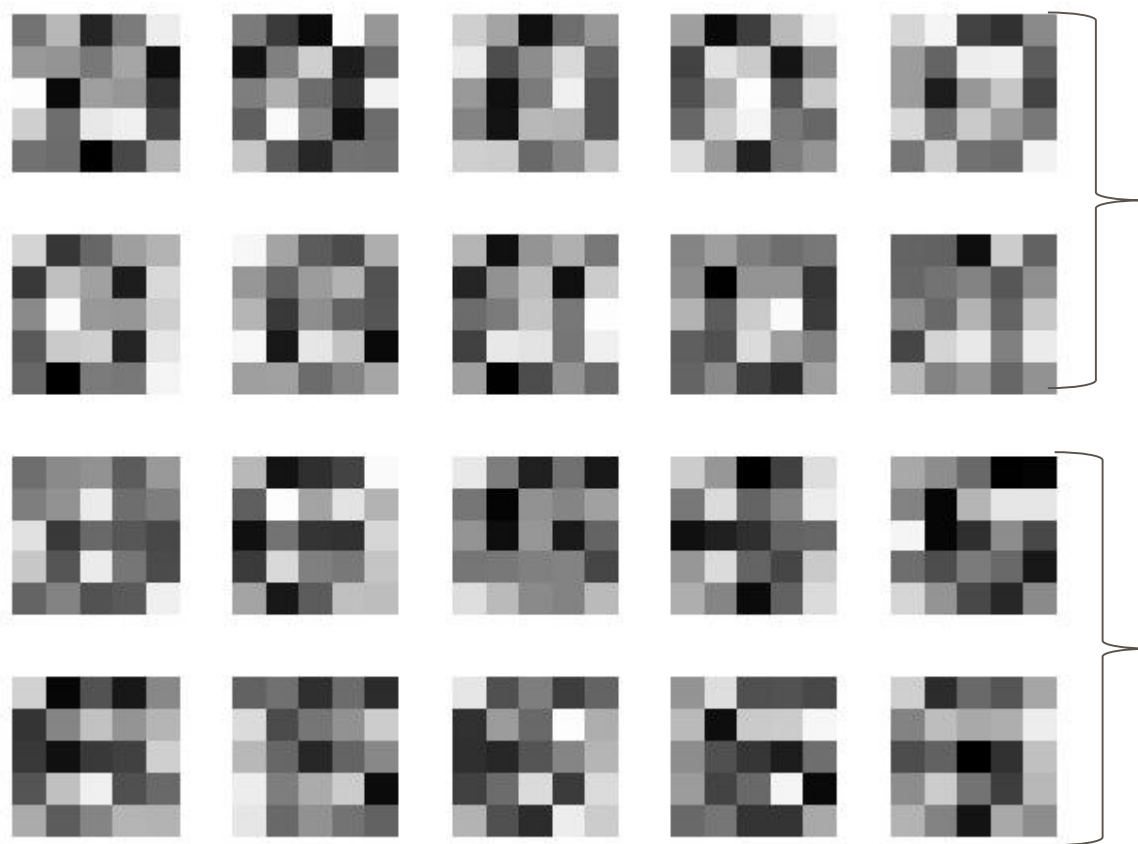


問題(3)

ロボットハンドに、ボールを投げさせて的にあてさせたい。ロボットハンド自身の重さや空気の抵抗や天候などのすべてを計測することは難しい。どうしたらいいだろうか？

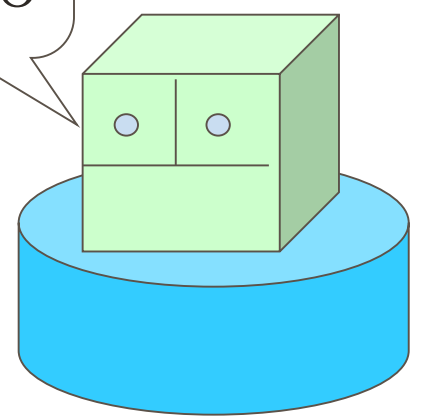
すずかけ台キャンパスに合計30台の自律ロボットを配備し警備に当たらせたい。どのようにしたらいいだろうか？

文字学習の例



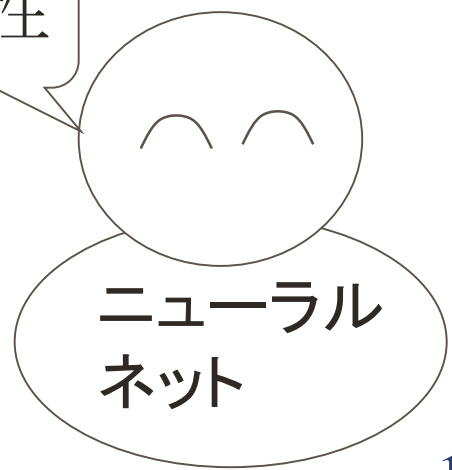
上が0
下が6

0

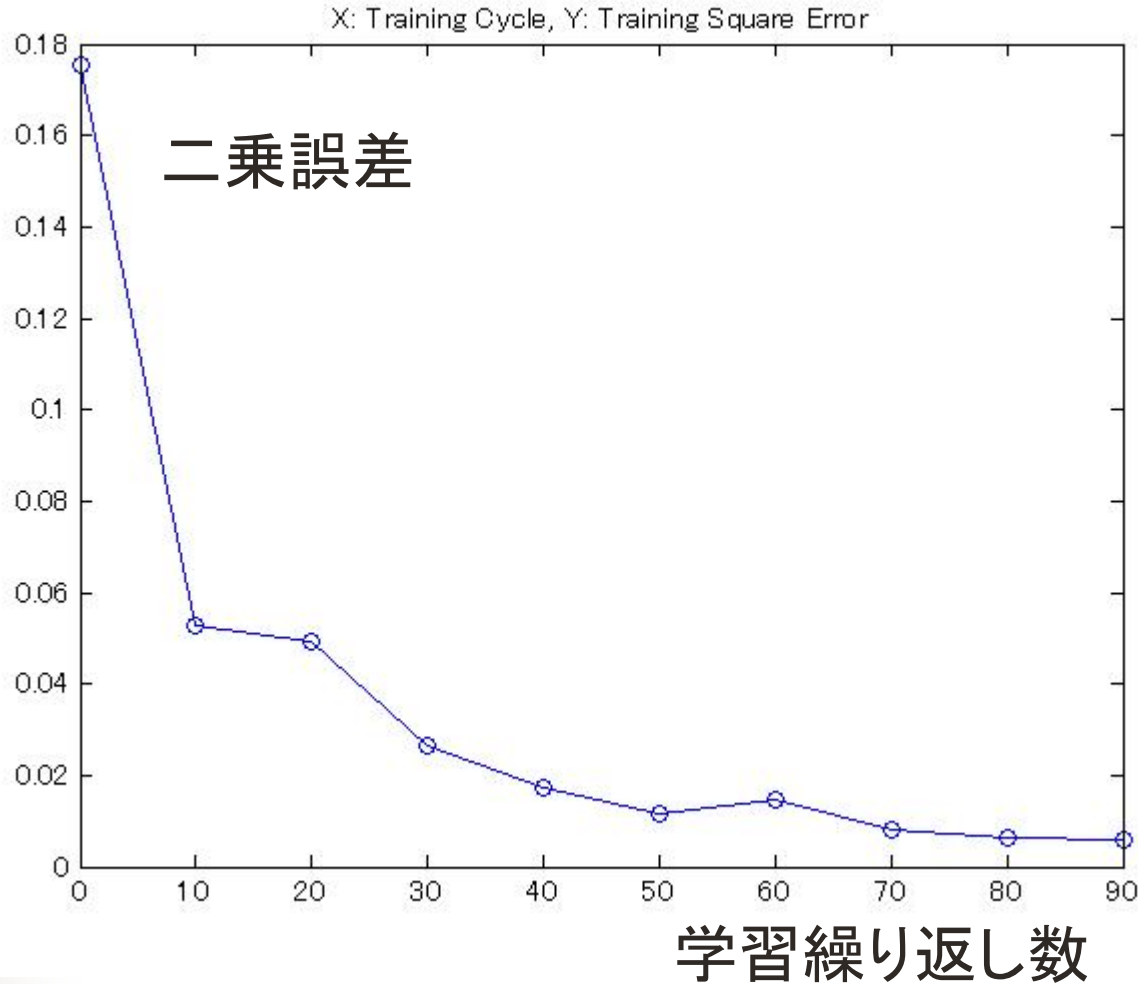


根性

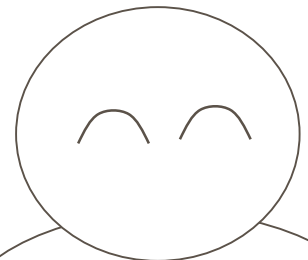
6



繰り返し学習



くりかえして
学習すると
だんだん
文字が読める
ようになる

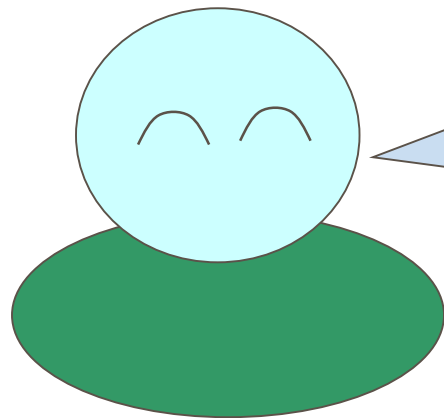


ニューラル
ネット

学習とは

応用：対象ごとに異なる

学習理論：どんな学習にも共通して現れる



僕には、意外に深い
数学的構造が
あるらしい・・・
・・・でも秘密。

この講義の目標は

「例からの学習」の基礎

パターン認識
システム制御
時系列予測

実験計画
データ解析
因果関係分析

確率論
数理統計
数理物理

例からの学習とは(1)

たくさんの例 x_1, x_2, x_3, \dots から
未来を精度よく予測したい

実データを扱うときに陥りやすい失敗

具体的なデータだけに固執して
その場その場の工夫を繰り返す

少しできたり、できなかったりの螺旋
・・・実質的に何も進歩していない



例からの学習(2)

たくさんの例 x_1, x_2, x_3, \dots

例を固定されたものと考えず

確率的にゆらぐ「何か」からの偶然の産物と見て

「何か」に名前をつけて「何か」をつきとめる

例からの学習(3)

$$X \longrightarrow X_1, X_2, X_3, \dots$$

例が増えると、〈例のもと〉がだんだんわかってくる

〈例のもと〉がわかると、未来が予測できる

できることとできないことが分かってくる

例からの学習(4)

例の元＝情報源 → 例₁, 例₂, 例₃, …

〈例のもと＝情報源〉を限界まで知るためには
どうするのが一番よいか？

例を見なくてはだめだが、例にこだわり過ぎてだめ。
「バランスがあること」を知ることが学問である。
バランスを定めている主役は何だろうか？

代数幾何学者・石井志保子先生の言葉
「小人, 例を忘れ, あるいは, 例に溺れる」



例からの学習(5)

学習理論の研究は世界中で盛んに行われている。

学習モデルのプログラムはたくさん公開されている。誰でも使える。

マスメディアで「人工知能」がセンセーショナルに宣伝されている。

研究開発に遅れた日本でも・・・始まった(?)。

しかし、まずは基礎から。ゼロからの出発です。

講義計画



ボルツマンマシン



深層学習

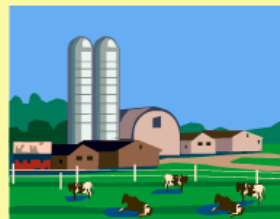


ニューラルネットワーク

教師なし学習



自己組織化



競合学習

教師あり学習



サポートベクタマシン