

# 医療情報と人工知 能・機械学習

京都大学EHR共同研究講座

特定講師 小林慎治

# Agenda

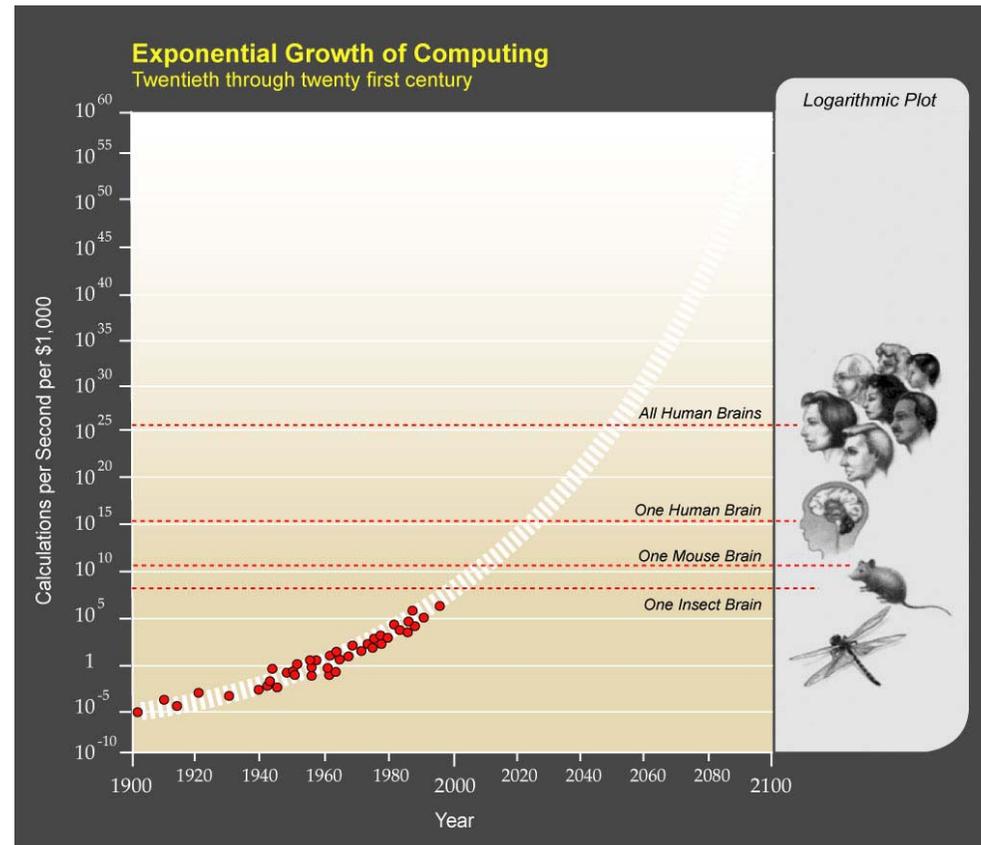
- AIとは何か
- 人工知能のあゆみと医療情報
- 現在の人工知能・機械学習の特徴
- 医療分野でのAIの現状と課題

# 注目される人工知能

- ボードゲーム
  - 将棋、囲碁、チェス、オセロ
- 画像解析・言語、音声解析
  - ネコ
  - 翻訳精度の向上
  - 自動運転
- ビジネス
  - 株価予測、在庫管理、クレーム対応
- 技術的進歩
  - ビッグデータ
  - ディープラーニング
    - 多層的ニューラルネットワーク

# シンギュラリティ (特異点)

- 人工知能が人間の知能を上回る点
  - 2040-2060年頃と予想されている
  - 悲観的、破滅的、楽観的、誇大妄想的なSF的予測がなされている
  - 多くの人が失業？



# Technology Will Replace Many Doctors, Lawyers, and Other Professionals

by [Richard Susskind](#) and [Daniel Susskind](#)

OCTOBER 11, 2016

[+](#) SAVE [↗](#) SHARE [26](#) COMMENT [HH](#) TEXT SIZE [🖨](#) PRINT **\$8.95** BUY COPIES



# AIとは何か

- 人工知能の定義
  - 「知性」「知能」の具体的な定義がない
  - したがって、人工知能そのものの定義も明確ではない
  - チューリングテスト
    - チャット形式で「人工知能」と会話させ人間が相手を人間と感じるかどうか
- 機能的定義
  - 推論
  - 探索
  - 知識表現

# 強いAI、弱いAI

## 強いAI

- 目標
  - 人間の代替、意識を持つ
- 演繹的
  - 論理ベース
- 汎用的
  - 自律的に回答をえることができる

## 弱いAI

- 目標
  - 特定の領域での問題解決
- 帰納的
  - 統計ベース
- 限定的
  - 個別の課題に対して過去の統計を元に解をだす。

# 医療分野でのAIへの期待

- 診断・治療精度の向上
  - 画像診断、病理組織診断
  - 手術用ロボット、治療方針決定支援
  - IBM Watsonによる診断事例
- 省力化、効率化
  - ワークフロー最適化
  - 文書作成支援
- 大規模データ処理
  - 「ビッグデータ」
  - ゲノムデータ

# 人工知能研究のあゆみ

- 第一次ブーム(1950-60年代)
- 第二次ブーム (1980年代)
- 第三次ブーム (2010年代)

# 第一次AIブーム (1950-60年代)

- 「推論と探索」
  - パターン認識、自動計画
- 成果
  - ニューラルネットワーク、パーセプトロン、Min-Max法、General Solver、Lisp、自然言語解釈、画像解析
  - ゲーム解法
- 終息した要因
  - コンピュータ性能の限界（組み合わせ爆発）

# 第二次AIブーム（1980年代）

- 論理プログラミング、遺伝的アルゴリズム
- 成果
  - エキスパートシステム(スペシャリストの判断を再現)
    - DECの発注の最適化
    - Mycin：伝染性疾患の診断
    - 心電図解析
  - Prolog
    - 看護師勤務表作成支援ツール
- 終息した要因
  - 資金不足、第五世代コンピュータ(シグマ計画)
  - そもそもスペシャリストの判断が論理的ではなく、標準化されてもいなかった。
  - 推論の元となるデータが十分にえられなかった。

# Racter, Mindscape, 1984



# 第三次AIブーム（2010年代）

- 「ディープラーニング」の実用化
  - ニューラルネットワークの進歩
    - たたみ込みニューラルネットワーク、再帰的ニューラルネットワーク
  - 計算機能の向上（GPUの応用）
  - 大規模データ収集と運用
- 成果
  - 折りたたみニューラルネットワーク(ディープラーニング)
  - ゲーム分野での人間への圧勝
  - 「シンギュラリティ」

# 5分でわかる機械学習

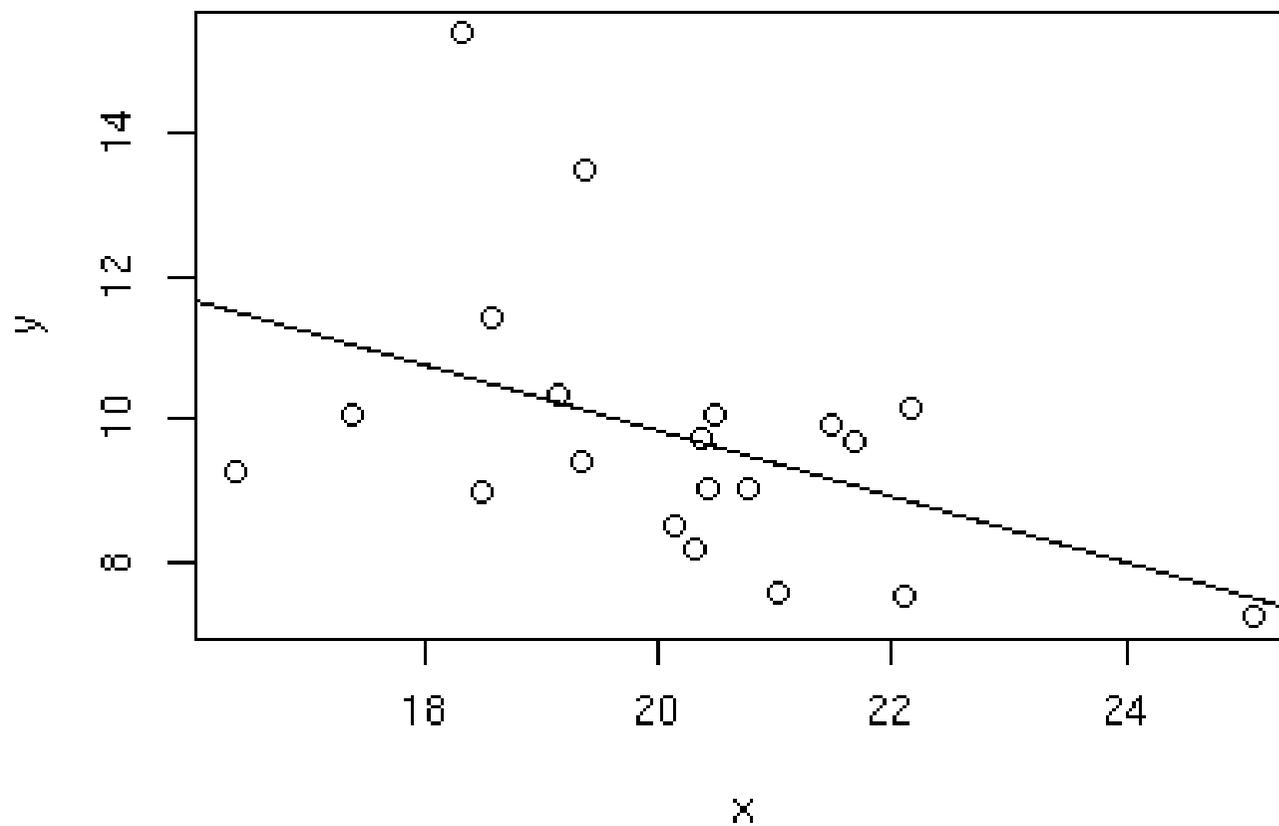
- 教師付き機械学習
  - あらかじめ学習用のデータが与えられる
  - 最適化問題
  - 分類問題
- 教師無し機械学習
  - 特徴を自動で抽出
    - クラスタリング

# 最適化問題（回帰分析）

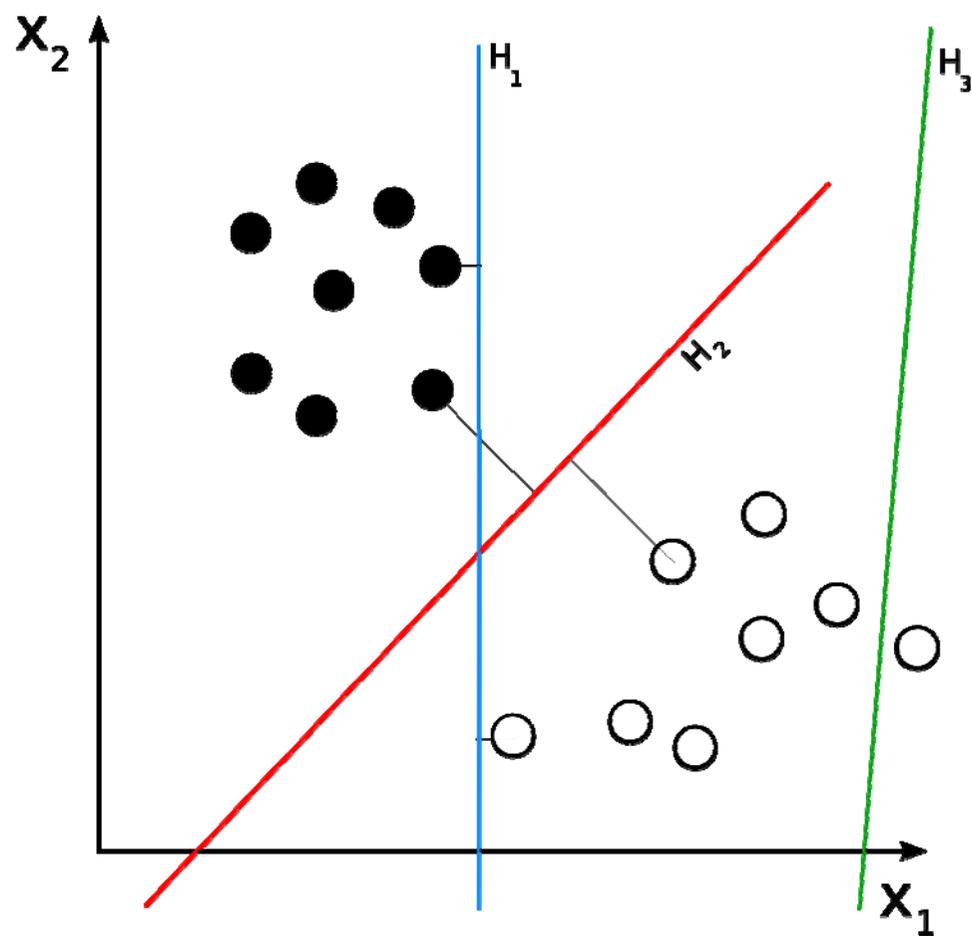
- データから特定の数値を予測する
  - 築年数、間取り、広さ、住所から家賃を予測する
  - 疾患リスク解析
  - 予後予測

Pt	性	年齢	WBC	...	OS
..					
1	M	13	3000		4.5
2	F	33	4500		3.2
3	F	18	2800		1.3
4	M	46	6000		5.4
5	M	21	1200		6.5
...	...	...	...	...	..

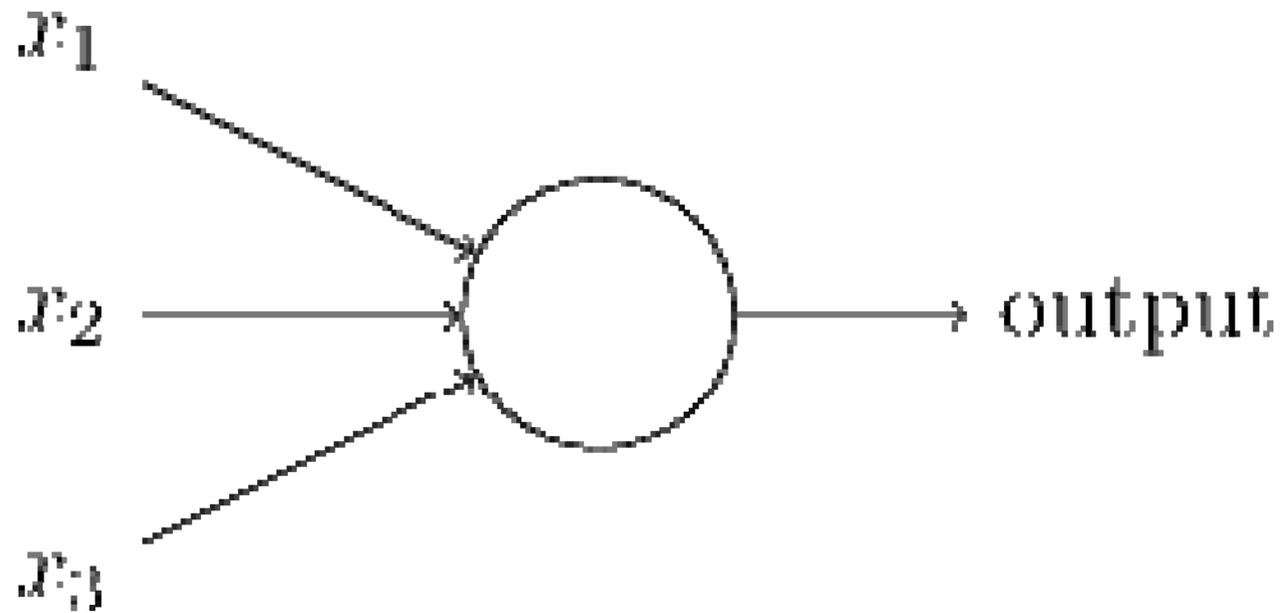
# 最適化関数



# 分類問題

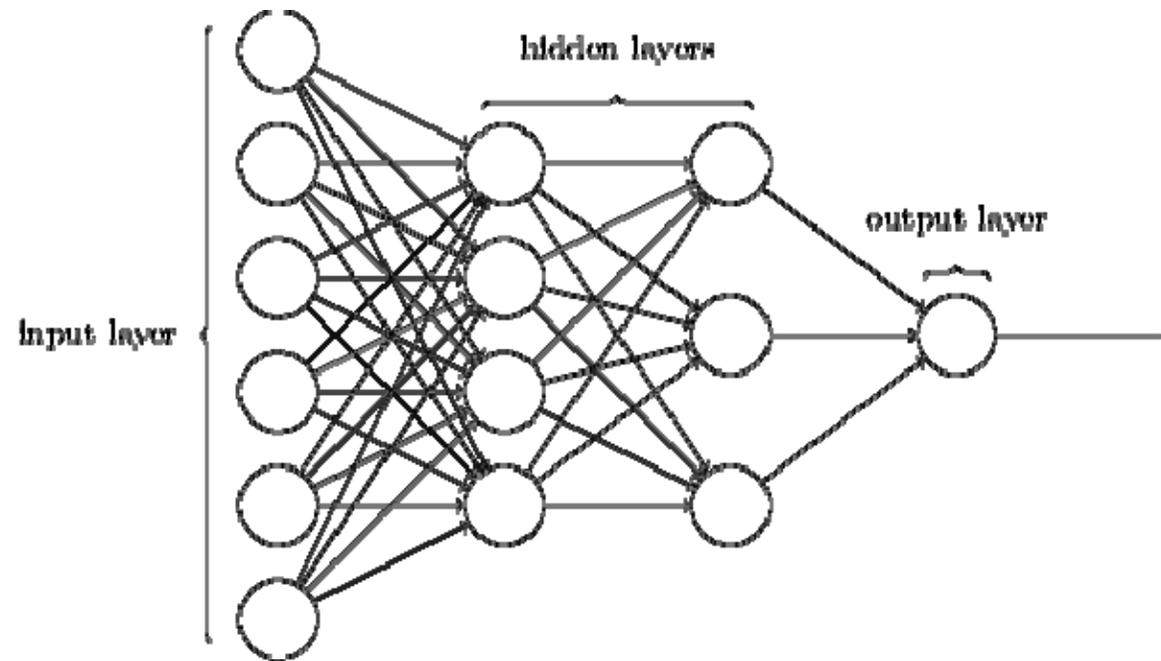


# パーセプトロン

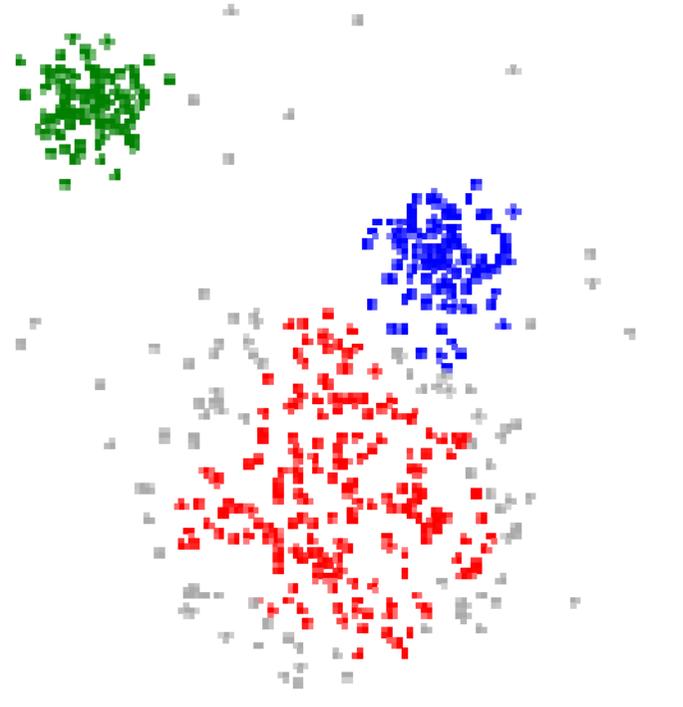


識別関数、サポートベクタマシン、単層ニューロン

# ニューラルネットワーク



# クラスタリング



# 「自動運転車」



# 既に実用化されている医療AI

- エキスパートシステム
  - Mycin
  - 看護師勤務表作成支援ツール
- 診断ツール
  - 心電図自動判定
  - 鳥越恵治郎「病名思い出しツール」
- 治療
  - AED
  - 放射線照射野設計

# 現在の開発中の医療AI

- IBM Watsonによる診断支援
  - Sloan Kettering癌研究所での研究成果
  - MD Anderson癌研究所での研究中止
- 画像診断、病理診断
  - 胸部Xpでの結節検出
  - マンモグラフィー診断
  - 病理での良悪鑑別、範囲の特定
  - Dermographyによる皮膚病変診断
- 自然言語解析
  - Twitterから感染症流行を察知
- ゲノム解析
  - SNPsと臨床プロファイルの関連づけ

# EHRとAI

- 「医療ビッグデータ」
  - 人類での初めての試み
  - Precision medicine
  - ベクトル化とパターン解析
- 診療支援
  - リスク分析、予後予測
  - 経営分析

# GPU Technology Conference 2017



# Example of Health Care Data

Example 1:

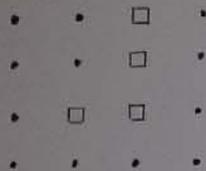
Example 2:



How are health care data different from the data from existing applications of deep learning?

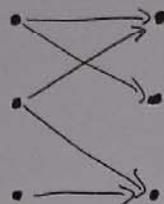
- Privacy, privacy!
- Heterogeneity
- Lots lots of missing data
- Big small data
- *Worst of all: doctors do not believe anything they cannot understand no matter how cool and how deep they are!!*

# Data Challenge 1 | Health Data is Dirty, Incomplete and Fuzzy



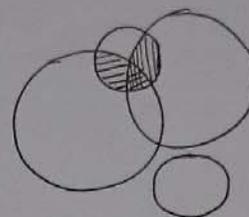
## Missing Data

No Lab Units or Ranges



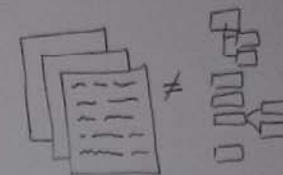
## N-M Mappings

N ICD 9  $\rightarrow$  M ICD 10



## Fuzzy & Overlapping Classifications

NDC  $\rightarrow$  RX



## Inconsistent Data

Clinical Notes = Unstructured

## Overview of Partners Data

### Radiology exam Data

Hospital	Historical records		Exams /year
	Exams	Images	
MGH	~10M	~1.8B	~0.8M
Partners	~20M	~3.6B	~1.6M
<b>Total</b>	<b>~30M</b>	<b>~5.4B</b>	<b>~2.4M</b>

### Other Data Items

RadMine (reports)	Biobank
RadStore (struct data)	RPDR
Cohort Store	Multiple small repositories

### EHR Records

Test type	#s
Lab Tests	~1.3B
#s for lab specimens	~136M
Micro specs	~16M
Path reports	~9M
Cardiology	~8M
Discharge summaries	~7M
Op notes	~3M
Endoscopy reports	~0.1M
<b>Total</b>	<b>~1.6B</b>

## Deep Patient: summary

---

- Pros

- ✓ Deep Patient enables to leverage EHRs towards improved patient representations
- ✓ This implementation uses a similar as simpler feature learning models
  - it can help to improve previous studies based on EHRs

- Cons

- ✓ representations are not interpretable
  - interpretability is a key only on predictive tasks
- ✓ time is not modeled



## Pulmonary Embolism Detection



Ziad Obermeyer, MD



↑ Risk PE



↓ Risk PE



## Lumbar Spine Segmentation



Stuart Pomerantz, MD



Radiology



Stenosis

# Real-Time Automated Segmentation of Challenging Suspicious Breast Masses

Viksit Kumar, Jeremy Webb, Adrianna Gregory, Max Denis, Mostafa Fatemi, Azra Alizad  
Ultrasound Research Laboratory, Department of Physiology and Biomedical Engineering  
Mayo Clinic, Rochester, MN

## Abstract

Computer aided diagnostic (CAD) tool for suspicious breast masses from the underlying algorithm is based on deep learning (CNN). The CAD algorithm runs on graphics processing units (GPUs) and Under a Mayo Clinic Institutional Review Board approval, a prospective study of the suspicious breast masses was conducted on 269 female patients who were undergoing suspicious breast masses, followed by core needle biopsy. The tool effectively segmented the masses with a 0.78 Sørensen-Dice coefficient. This allows the algorithm to have segmentation, which can be used for clinical applications. The algorithm has the potential for real-time monitoring, where ultrasound is

## Objective

It is estimated to be the most common cancer among women in 2016<sup>1</sup>. Breast cancer is typically detected by mammography followed by ultrasound. The detection of a suspicious mass on ultrasound and data system (BIRAD) scale is a purely visual task. The basis for

## Cohort

This research uses 490 breast ultrasound images, of one or more viewing planes, obtained from 269 patients gathered by clinically-certified experienced sonographers. Images were obtained from patients seeking care at the Mayo Clinic, with Institutional Review Board approval and proper consent. 30 images are of normal breast tissue, and the remaining 450 images contain breast masses. Approximately half of the masses are biopsy proven benign and half are biopsy proven malignant. Images were segmented by hand. The accuracy of these hand segmented images was confirmed by a trained sonographer with 28 years of experience. As shown in figure 3, most of the images are from BIRAD 4 patients. These cases are especially difficult to assess risk for as they don't have features specific to a pathology.



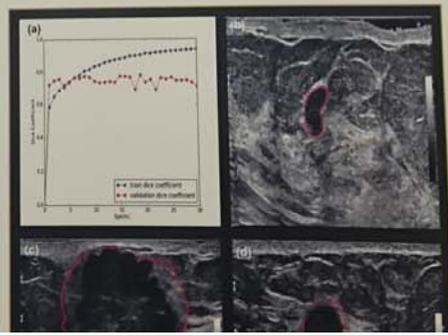
Figure 3. Bar plot showing patient distribution by BIRAD rating. A BIRAD score of 1, 2 and 3 will have characteristics of healthy tissue or appear benign. A BIRAD score of 4 is inconclusive, does not possess clear features. A BIRAD score of 5 has features suggesting malignancy. BIRAD 6 are cases confirmed malignant by biopsy. BIRAD 0 is incomplete and requires more imaging.

## Methods

We utilized the CNN based U-net model proven effective at medical

## Results

The Dice coefficient for the training and validation set is presented in figure 5(d) as a function of increasing number of iterations. The close overlap of the segmented and the predicted segmentation showcases the efficacy of the algorithm. The ability to train the algorithm to a Dice coefficient of nearly 1.0 suggests that the algorithm is capable of learning all the textural features required to segment the mass. Overfitting can be estimated by tracking the degree of separation of the training and validation sets. And can be seen starting at epoch nine. When overfitting occurs, the model is starting to learn features specific to the training set and will not generalize to new data. Overfitting is the greatest obstacle when using deep learning.



## Conclusions

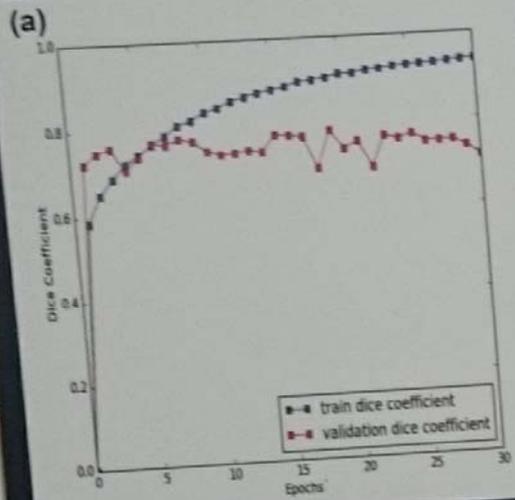
We have developed an algorithm which can automatically segment suspicious masses in real time ultrasound exams without the requirement of an initial seed. This has benefits in reducing exam time and aiding training of amateur sonographers thereby freeing up time for medical specialists for more complex tasks. In remote areas without specialized modalities (mammography) or radiologists, such an algorithm can aid in breast cancer screening.

## Future Work

- Future work includes hyper parameter optimization of the deep learning model.
- Expanding the data set. Larger datasets will improve segmentation accuracy and improves the ability to generalize to new images.
- Using a modified version of this model to develop a combined classification-segmentation algorithm. This will allow for the inclusion of biopsy results i.e. ground truth into training of the algorithm.

## Performance

ew data. Overfitting



ground B-mode images of breast masses (manual)

# 医療分野でのAIの問題点

- 学習内容が過去のデータの集積である
  - つまり、「後方視的」
  - 相関関係を示すことはできるが因果関係を示すことができない
  - 限定的な課題では威力を発揮することができるが、汎用的な問題を解決することはできない
- 言語処理能力がまだ不足している
  - 医学論文を読み解いて、内容を臨床データに当てはめて評価することができない
  - したがって、人間がルールセットを記述して機械に教えてあげる必要がある。
- 時系列処理が苦手である
  - データセットをそろえることが難しい
  - 次元削減やLSTMなどのテクニックが必要

THIS IS YOUR MACHINE LEARNING SYSTEM?

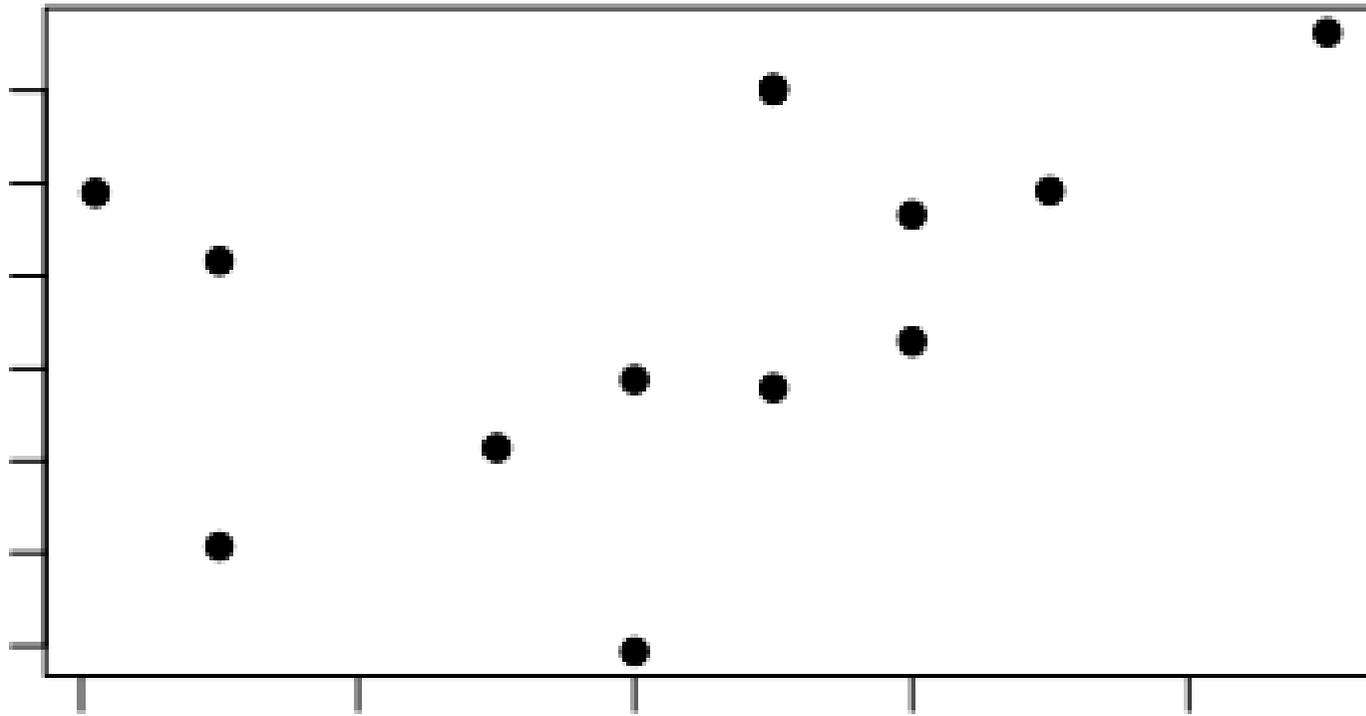
YUP! YOU POUR THE DATA INTO THIS BIG PILE OF LINEAR ALGEBRA, THEN COLLECT THE ANSWERS ON THE OTHER SIDE.

WHAT IF THE ANSWERS ARE WRONG?

JUST STIR THE PILE UNTIL THEY START LOOKING RIGHT.



# 機械学習のチューニング



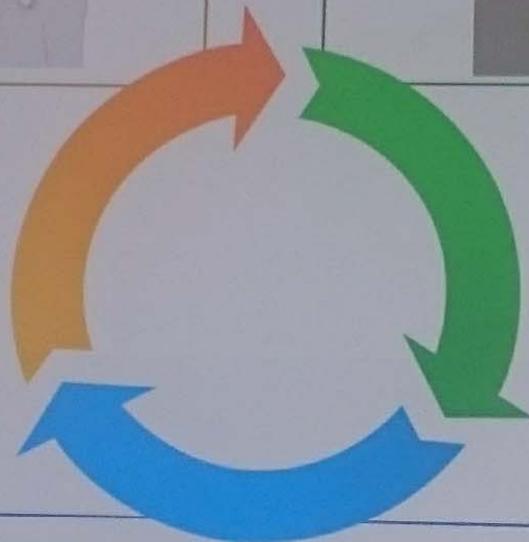
### 1. Data

*Unprecedented, long-term, multi-dimensional data*



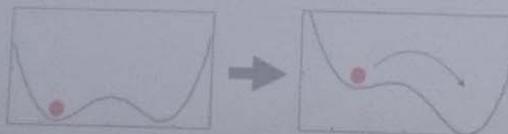
### 2. Deep Learning

*Powerful algorithms for large-scale data analysis*



### 3. Theory

*Novel insights from Dynamical Systems Theory*



# 医療でのシンギュラリティはくるか？

- 限定的には到達
  - 一部のガイドラインベースでの診断支援、プロセス最適化
- 医学の進歩は人智を超えつつある
  - 専門医でも自分の分野の論文やガイドラインをすべて網羅することが難しくなりつつある
- 医師は不要となるまで機械が進歩できるか
  - 電子化された診療記録から診断を導くことはある程度可能
  - そもそもの診療記録に間違いがある
  - 患者から得た情報を電子的に記録するまでに医師の主観や診断プロセスが入っている
  - 医師の代わりに問診、診察、診療記録を作成するAIは、「強いAI」であり、現在の技術では不可能

# まとめ

- 人工知能、機械学習について
  - これまでの流れと現在のトピック
  - 教師付き機械学習、回帰分析、パーセプトロン、ニューラルネットワーク
  - 教師なし機械学習、クラスター解析
- 医療分野での人工知能、機械学習