

2025.1.30

ベイジアンネットワークを用いた 医療データの因果探索



張山昌論 さん
東北大学大学院
情報科学研究科

人にやさしい計算システムの開発

① システム・ソフトウェア
× ② ハードウェア

テーマ1.

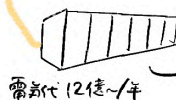
知能システムの
研究開発

コミュニケーション手段
のためのAI



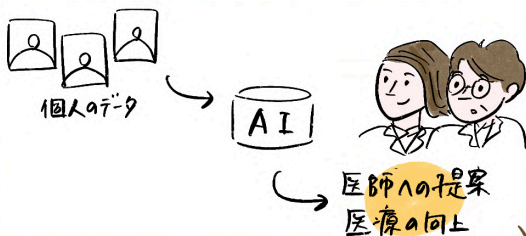
テーマ2.

AI/ビッグデータ解析
のための専用コンピュータ開発



スーパコンピュータを
超える手軽に!

医療情報処理システム



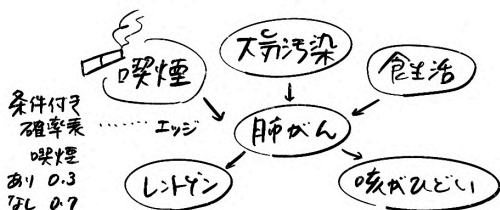
データ解析に用いる

AIの望ましい性質

- モデルの理解NO
- ランダム
- ブラックボックス
- 大量データ必要!
- タイプランニング
- 交換値NG
- ベイジアンネットワーク
- 少量データOK
- 交換値OK
- モデル理解容易

ベイジアンネットワーク(BN)

事象間の確率的な依存関係を
グラフでモデリング



EX.1

腹腔鏡下での手術が困難な症例



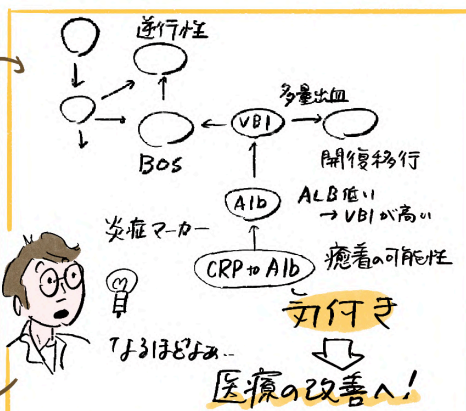
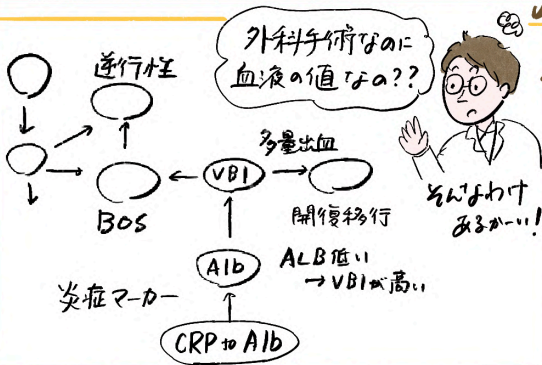
持病量=37

Q データをいれると
アルゴリズムに従って
BNは
ユニークに
設定されるの?

A 基本的にはそう!

1. アルゴリズム
2. 評価関数

2つの組み合わせで
ネットワークの形は
変わります!



ベイジアンネットワーク

全体的な動きが得られる

連続変数の扱いがむずかしい

適切な閾値を設けて離散化する

ネットワークが変化するので
設定がむずかしい

代表的な変数の離散化手法

- 等間隔
- 等頻度
- K-means法
- カスタム法
- 意味考案あり
- 密度
- 経験則・知識が重要

2025.1.30

ベイジアンネットワークを用いた 医療データの因果探索

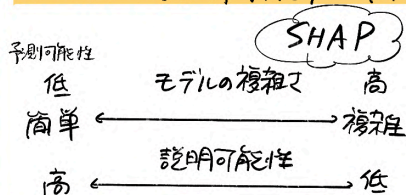
説明可能AIを活用した
高信頼ベイジアンネットワーク構築

AI(SHAP) → BNの閾値
真似る

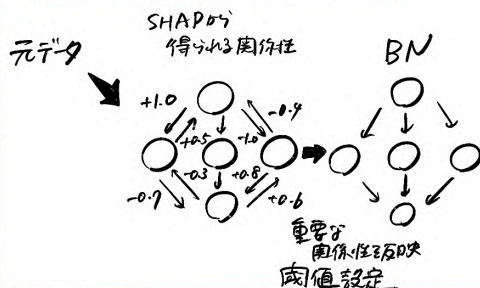
医療系のAIモデルにおける
説明可能AIの重要性



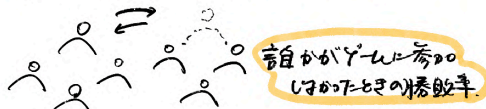
AIモデルと説明可能性の関係



SHAPを活用したBN構築

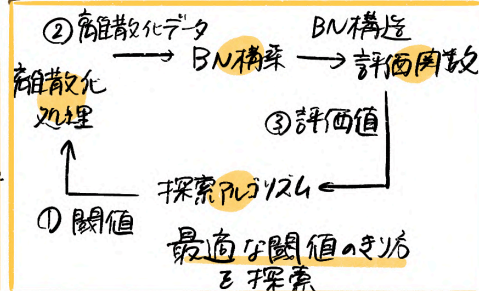
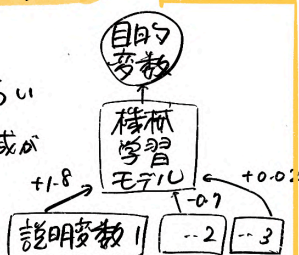


そもそもSHAPとは?
ゲーム理論における評価関数



誰か = 変数の貢献度・重要度

- SHAP値が大きいほど重要度が高い
- 説明変数と目的変数の増減がSHAPの符号と一致



2型糖尿病

人間ドックと生活習慣のアンケートを統合したデータ
5,879名、20変数 ちゃんと答えがそろー

生活習慣

血液成分

BMI

分散化



名分散化手法との比較

	等間隔	基準値	提案
BMI	38.7	25	23.7
TG	214	150	259
HDL	83	40	54
LDL	174	120	127
評価値	7.74	15.16	16.60

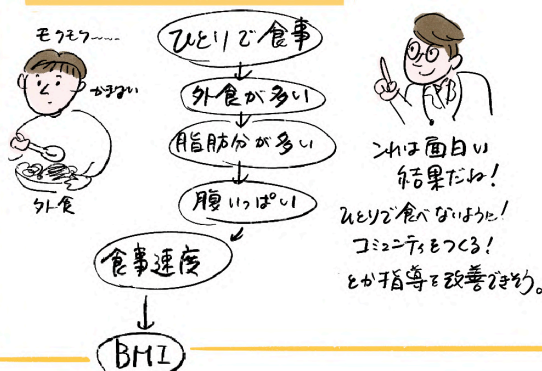
閾値の切り方に
見通しが立ってきた!

② 閾値の上か下か
2値?

A 2値で可。
3値に可ともう一歩さらに
意味のある値があるとき。
複数の影響を考慮
すると可。



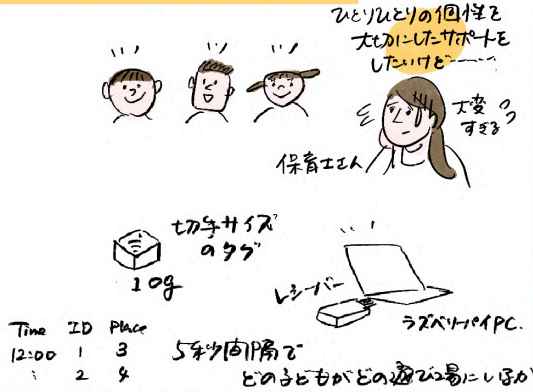
ネットワークの構造比較



2025.1.30

ベイジアンネットワークを用いた 医療データの因果探索

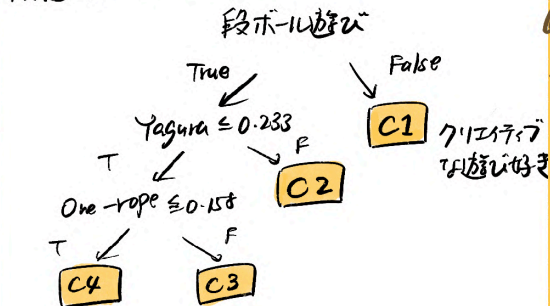
子どもの発達を測るシステム



プレイパーク 山形県ほろ市

Time	Child 1	2	3
12:33:30	Shichirin	Singrope	
12:33:40			
12:33:50			
	Shichirin	Cardboard	滞在時間 グループ分け
child 1	0.10	0.50	C1
2	0.30	0.03	C2
3	0.05	0.80	C3

決定木にかけて...



より良い遊びを提供する /

向山子ども園と連携して、東北の家の
2月から、実馬舎を開始。

会話特徴を用いたAIモデル

SHAPを用いた解釈

ASDでない ASD

発話

フレー

文脈類似度平均

文脈類似度分散

色んな

話題に比べ

形容詞

感覚的

?

同じようなことを話している

?

?

?

?

?

?

?

?

?

?

?

?

?

?

?

Q 類似度の平均と分散とは？
変り方のこと？

A 特に平均とは？
は学生と議論中...

コミュニケーション支援のためのAI技術



学習障害の発見を目指した時系列データ



Q バイジアンネットワークについて

「エッジに確率がのっている」

ひとり食べる 閾値「0食に0回ひとり」

外食が多い 0%の確率で次にいく

確率が2つのっているということ??

A 閾値はポイントで段階くらいに1%かたまりで決めている
セカンダリな提案手法で最適化

セカンダリによって

前半のグループと後半のグループが異なり、異なる関係性があります。

エッジ(関係性)が理れるかどうか...

条件に

確率がある。

BMIが0.7だったUAは...

2025.1.30

ベイジアンネットワークを用いた 医療データの因果探索



Q. どちらがもっとも良い良くないかの
評価基準を入れるのは?
あるいは答えが分かっているものから
生成されたデータで実験する?

A.

因果関係を設定したデータ
をもとに... みると思います!



Q. BNは因果関係を表していると思ってもいい?

A.

一応、因果と言っている
のでは...
単純に相関というより、
関係のないデータは除いているので...



Q. BNは向きが分いたデータ
ですかね?

A.

最初は向きのないシステムで
次にどちら向きかがわかります。



Q. SHAPについて...
これは全結合になる??

A.

マトリクス、すべての変数から
すべての変数への
関係性。

閾値をかえることで
見える関係性が変わってくる。

- ① SHAPで強い関係性が得られる
- ② ブートストラップによって、ネットワークを作ってから...
信頼性の高いエッジを探索

BNはルールが構築できるかもしれない。



Q. BNは因果関係だとして...
SHAPがやっていること
因果関係として解釈できるものに
帰納と因果という点か。

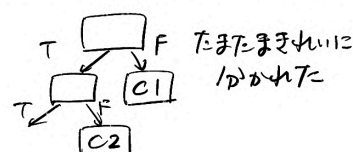
A.

SHAPは...
相関をみることで可能性が高い。

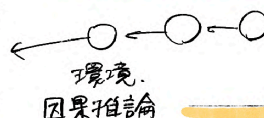


Q. 子どものみそびのデータについて。
分類しているから因果ではないのか。
分類した結果の要因分析を
していくのか??

A. 決定木



そうである。属性、兄弟構成etcとグループを見ると
因果がもう少し見えてくるかも。



2025.1.30

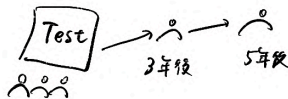
ベイジアンネットワークを用いた 医療データの因果探索

Q 子どもにあずめる
スポーツを変えられる？

息子に
言うかは
別ですが
.....



A 追跡調査や全国でできた
可能かと思いますが！
実際、米国の水泳スポーツでは
そうしているみたいですよ。



Q AIモデルについて

- ・決定木が類基準
- ・BN - 因果関係

一階層のもの

A BNとAZモデルは
同じ構造ではないので
一概にはいえない。

Q どんな閾値においても
残るネットワーク、残らないネットワー
クもみることって...

因果関係の強弱は分かる??

A 閾値の切り方・切れ目
みえる関係性が変わる...
2値・3値でも...変わってくる。

Q お医者さんへの説得性
はどのように影響そうしてるか？
モデルに依存している？



A 本質的にはモデルは関係ないけど、
ひとつのモデルだけでなく、
複数のモデルを出す！



このお医者さん
案件には
どんなモデル
がいいかな...



Q 検証的なものはできたりしますか？

A できたら面白いですね！

2025.1.30

ベイジアンネットワークを用いた 医療データの因果探索

Q 「因果みたいなものがみえてきた」
重みを感じました...

A 感覚的に、「因果の流れ」が整っている
納得性がある、専門家との相談

Q SHAPが説明可能性を
あげたということですか？
我々は因果と関係のところが原因・
結果が分かると説明可能性が
あがると思う...

A そうですね。
・目的変数に対してどの説明変数が重要な
・どの方向に作用するか。

LINGAMを使うとはありますか？

あります！基本的に連続変数なので若干使いにくい...

たとえば、夏と冬の電気使用料・時間が違うときに
それぞれ因果グラフ・条件分けすると隣接行列が変わる...

行列同士の距離で定義すれば

異なってくる因果関係の通し・逆も測れるのでは？

という研究をしています...

閾値によってグラフが変わると

「条件ごとに因果関係が変わる」が似ていると感じます！

似ていると思いますね。

条件分けの正当性が分かってくるといいですね。