

心臓病罹患に関する分析

Team3 : 長田 茅野 河合

概要

✓ 調査

- ドメイン知識から重要なカラムを選出

✓ 分析

- 選出したカラムを可視化
- 事前に影響力を予想・考察

✓ 結論

- 実際にモデルに適用し影響力を確認
- 結果から考察

1.

✓ 調査

- ドメイン知識から重要なカラムを選出

✓ 分析

- 選出したカラムを可視化
- 事前に影響力を予想・考察

✓ 結論

- 実際にモデルに適用し影響力を確認
- 結果から考察

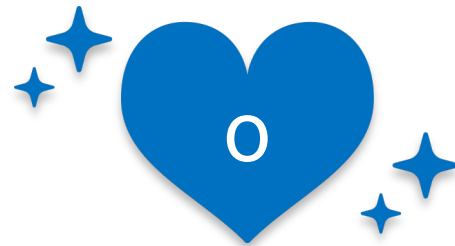
心臓病について

■ そもそも「心臓病」とは

- ・ 心疾患に関する症状の総称
- ・ 心不全, 虚血性心疾患, 心臓弁膜症, 心筋症, 不整脈, 先天性心疾患 等



「心臓病」という意味が広義的であることから
さまざまな要素が関係してくるのでは？



or



心臓病について

■ 心臓病の主な原因

糖尿病



FastingBS

- ✓ 概 要 : 空腹時血糖値
- ✓ 単 位 : mg/dL
- ✓ 正常値 : 70~110 mg/dL
- ✓ 備 考 : 120mg/dL以上を1, それ以外を0

高脂血症



Cholesterol

- ✓ 概 要 : 血清総コレステロール
- ✓ 単 位 : mg/dL
- ✓ 正常値 : 120~220 mg/dL

高血圧

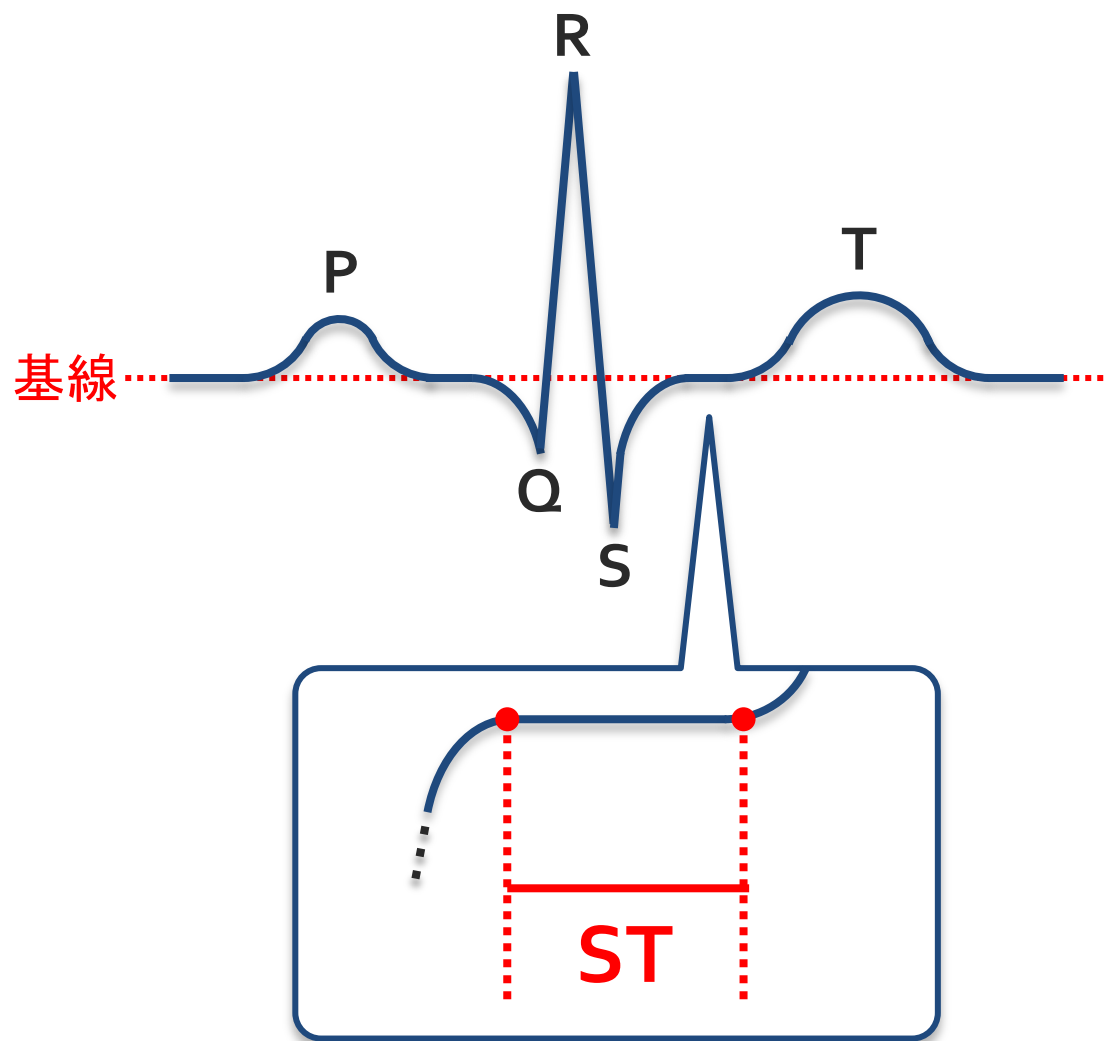


RestingBP

- ✓ 概 要 : 安静時最大血圧
- ✓ 単 位 : mmHg
- ✓ 正常値 : 一般に140mmHg未満
糖尿病患者は130mmHg未満

心臓病について

■ 心臓病の診断方法



✓ ST上昇

- ・ 概 要 : ST部分が上がった状態
- ・ 診 断 : 心筋梗塞、心筋炎、ブルガダ症候群
- ・ 備 考 : 病気がなくても現れることがある

✓ ST低下

- ・ 概 要 : ST部分が下がった状態
- ・ 診 断 : 心筋虚血、心肥大
- ・ 備 考 : 傾きによりUp, U字, Flat, Downに分類

✓ ST正常

- ・ 概 要 : ST部分が基線に近い状態
- ・ 正常値 : $-0.5 \sim 2.0$ mm

心臓病について

■ 心臓病の診断方法

ST異常



Oldpeak



ST-Slope

- ✓ 概 要 : ST波から基線までの距離
- ✓ 単 位 : mm
- ✓ 正常値 : $-0.5 \sim 2.0$ mm

- ✓ 概 要 : ST波の傾き
- ✓ 単 位 : -

2.

✓ 調査

- ドメイン知識から重要なカラムを選出

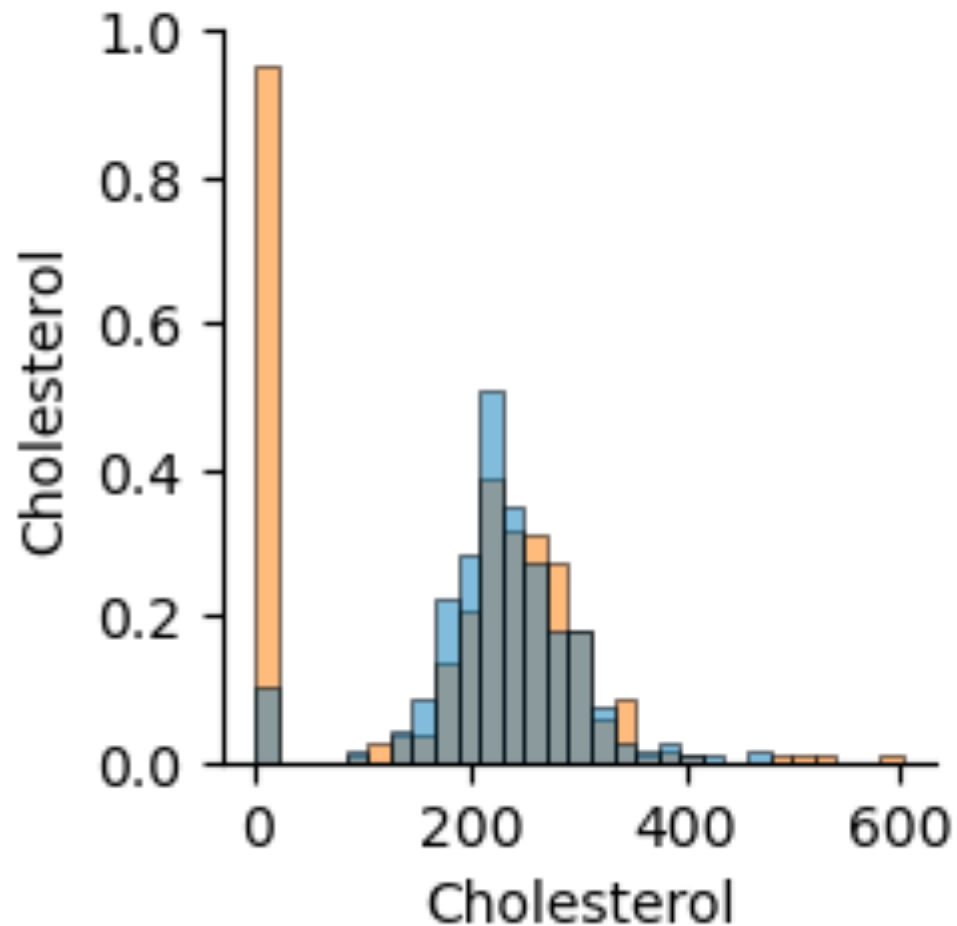
✓ 分析

- 選出したカラムを可視化
- 事前に影響力を予想・考察

✓ 結論

- 実際にモデルに適用し影響力を確認
- 結果から考察

Cholesterolカラムについて



✓ Cholesterolとは

- 概 要：血清総コレステロール
- 単 位：mg/dL
- 正常値：120～220 mg/dL



- ✓ 全体的に正常値と比べて値が高い
- ✓ 0は欠損値と判断

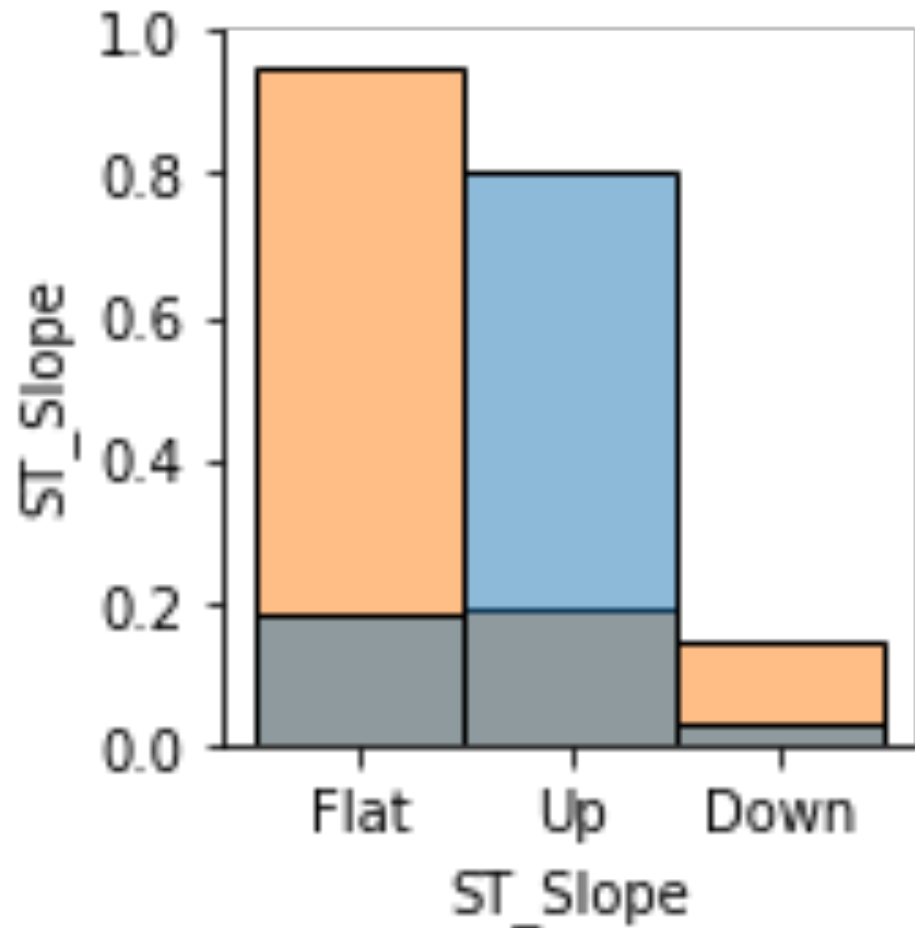
ST-Slopeカラムについて



Blue
HeartDisease:0



Orange
HeartDisease:1



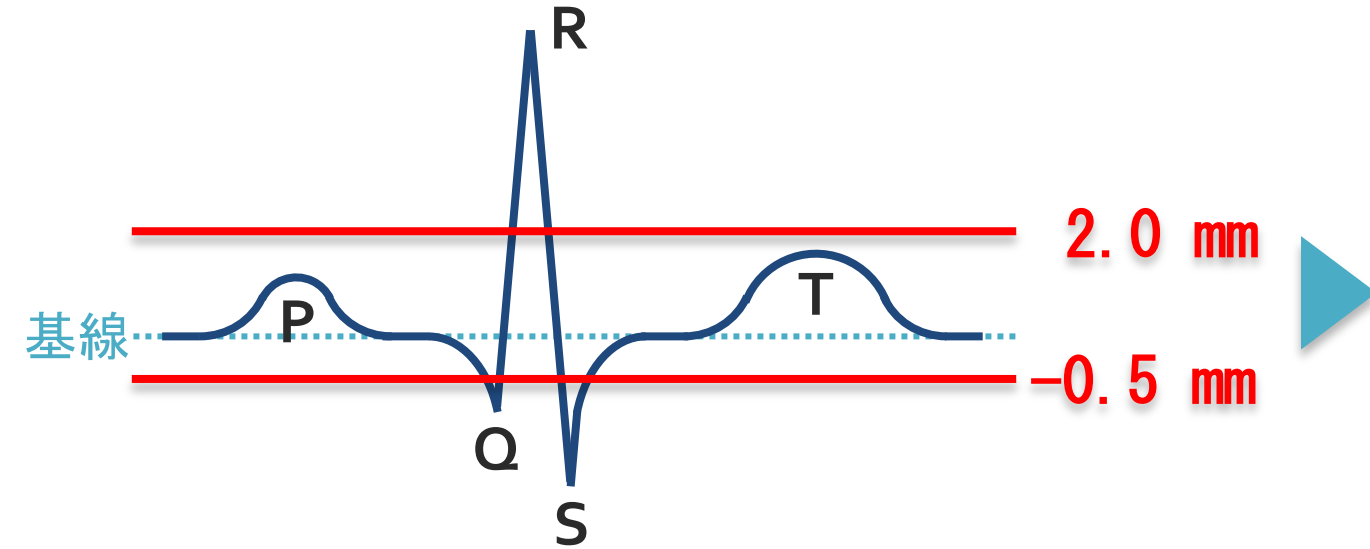
✓ ST-Slopeについて

- 概 要 : ST波の傾き
- 単 位 : -



- ✓ 「心臓病になりやすい」予測
→ Flatが影響するのでは
- ✓ 「心臓病になりにくい」予測
→ Upが影響するのでは

ST-Slope, Oldpeakカラムについて



- ✓ Abnormal + : 2.0mmより大きい
- ✓ Normal : -0.5mm以上2.0mm以下
- ✓ Abnormal - : -0.5mmより小さい



ST-Slope

ST-Slope, Oldpeakカラムについて

Heart Disease Old-ST	Abnormal+			Normal			Abnormal-		
	Up	Flat	Down	Up	Flat	Down	Up	Flat	Down
0	—	2	1	223	49	8	—	—	—
1	7	40	19	46	222	19	1	3	7



- ✓ 「心臓病になりにくい」予測
→ 「Oldpeak : 正常 | ST-Slope : Up」が影響するのでは
- ✓ 「心臓病になりにやすい」予測
→ 「Oldpeak : 正常 | ST-Slope : Flat」が影響するのでは

3.

✓ 調査

- ドメイン知識から重要なカラムを選出

✓ 分析

- 選出したカラムを可視化
- 事前に影響力を予想・考察

✓ 結論

- 実際にモデルに適用し影響力を確認
- 結果から考察

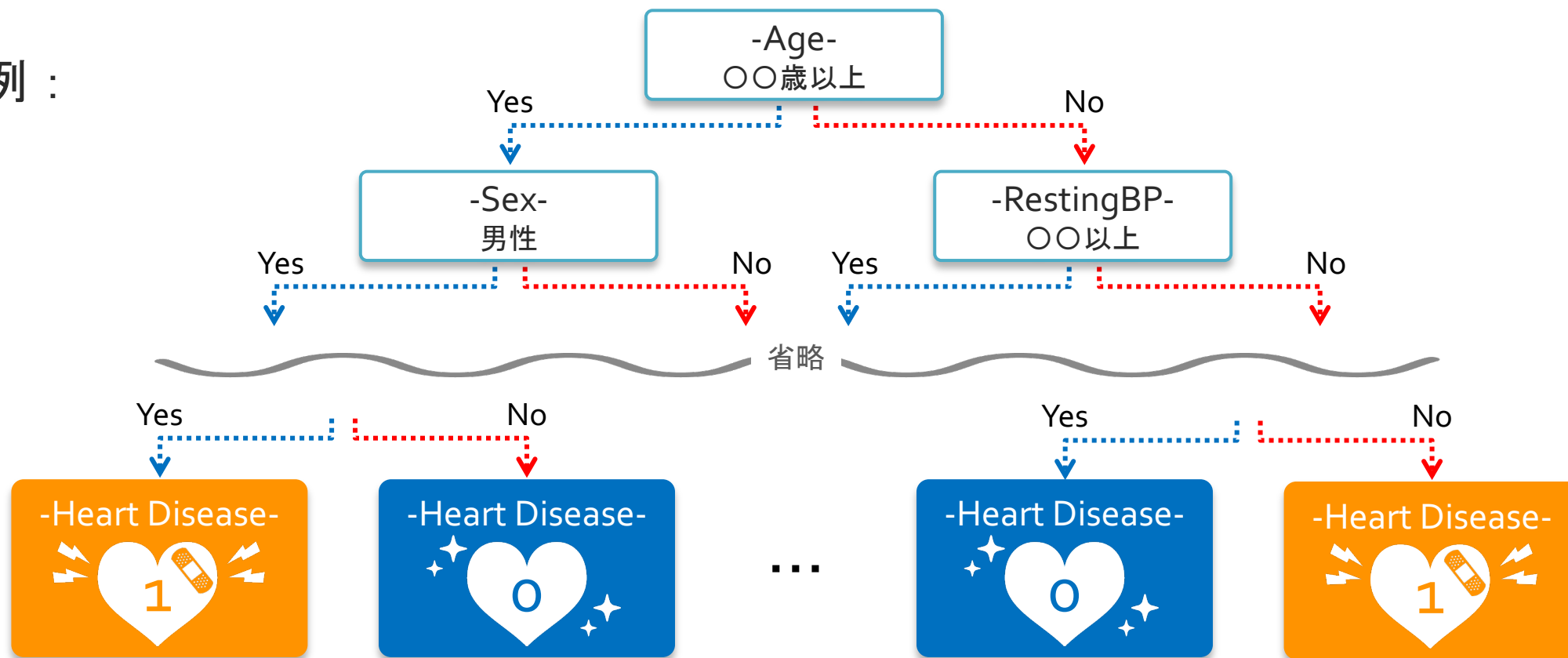
学習方法

■ 決定木

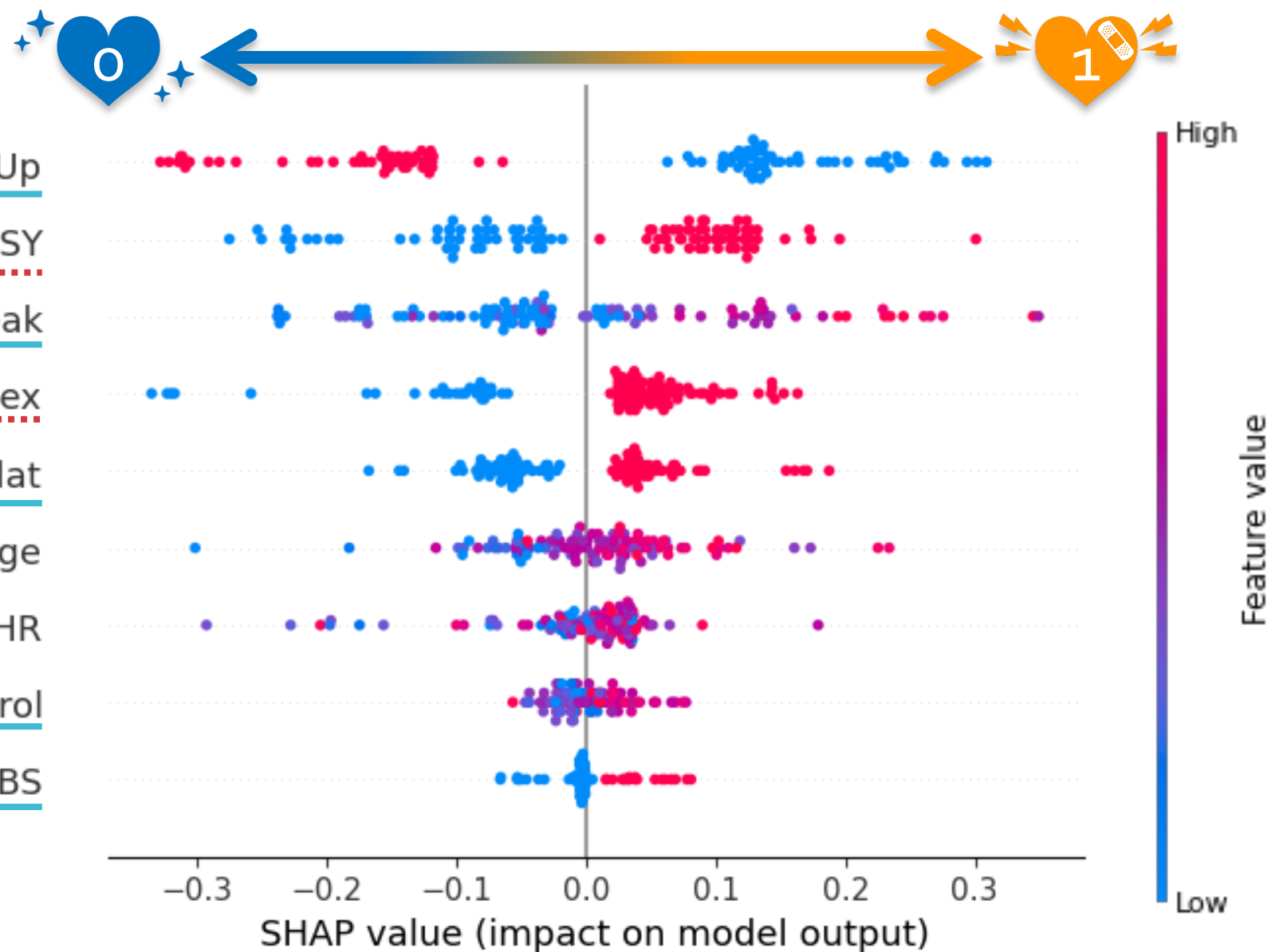
樹形図を用いて分類・回帰を行う機械学習手法

→ 複数組み合わせで、より誤差を小さくするよう拡張したXGBoostを採用

例：



学習の結果



✓ 注目カラム

✓ 新事実


なりやすい人まとめ

■ 事前に調査・分析した特徴

ST-Slope
ST波の傾き

- 
1. Upでない
 2. Flatである


Oldpeak
ST波の離れ具合

- 
3. 高い

Cholesterol
血清総コレステロール値

- 
4. 高い

FastingBS
空腹時血糖値


- 
5. 1である
糖尿病の疑い

■ 未調査の特徴

ChestPainType
胸痛の種類

- 
6. ASY
無症状である

Sex
性別

- 
7. 1である

カラムについて考察

6. ChestPainType : ASY
→ 胸痛に関して無症状

ChestPainType_ASY



■ 再分析

- ChestPainType : ASY | ST-Slope : Flat

HeartDisease	
1	0
203	19



- ✓ Flatと合わせて抽出すると Heart Diseaseの人数が高い
- ✓ Flatであることに帰着する
- ✓ 単体で影響力が強いと言えない

カラムについて考察

7. Sex : 1 → 性別が1



■ 再分析

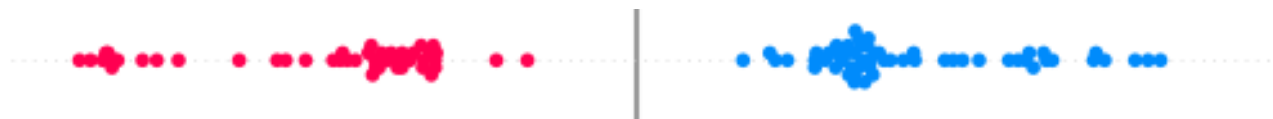
Heart Disease \ Sex	1	0
1	324	35
0	181	102

- ✓ 元々割合が高い
- ✓ 性別：0が極端に「心臓病でない」ことに影響
- ✓ 単体で影響力が強いと言えない

カラムについて考察

1. ST-Slope-Up : 0
→ STの傾きがFlatかDownである

ST_Slope_Up



4. ST-Slope-Flat : 1
→ STの傾きがFlatである

ST_Slope_Flat



-
- ▼
- ✓ 「なりやすい」という点に関してST-SlopeがFlatであることに帰着する
 - ✓ 「Oldpeak : 正常 | ST-Slope : Flat」が強く影響していると考察

カラムについて考察

- 「Oldpeak : 正常 | ST-Slope : Flat」について
 1. ドメイン知識が異なる
 - 「ST波の傾き」という解釈が違うのでは
 - FastingBSやRestingBPのように、特定の条件下で測る指標では
 2. 他に重要な要因がある
 - ST異常について、2つのカラムで考えたが、他に要因があるのでは
 3. データに偏りもしくは誤りがある



- ✓ 再考察にするには、**サンプルサイズが小さい**
- ✓ データの偏りが見られるため、今回のデータから考察しにくい
(Cholesterolが全体的に高い傾向)
- ✓ ラベルに誤りがないか確認すべき

なりやすい人まとめ

■ 影響があると考えられる特徴

Oldpeak
ST波の離れ具合

▶ 1. 高い

Cholesterol
血清総コレステロール値

▶ 2. 高い

FastingBS
空腹時血糖値

▶ 3. 1である
糖尿病の疑い

■ 再検討すべき特徴

ChestPainType
胸痛の種類

▶ 4. ASY
無症状

ST-Slope
ST波の傾き

▶ 5. Flatである

Sex
性別

▶ 6. 1である

- ✓ 妥当な**サンプルサイズ**の検討
- ✓ **抽出方法**についての再考