

主催：株式会社福岡ソフトリサーチパーク

後援：福岡市、福岡経済同友会、一般社団法人九州ニュービジネス協議会

公益財団法人九州先端科学技術研究所

# AI応用 活用上のポイント

実装の事例から

Aswright Logic

代表

畑 雅之

HATA MASAYUKI

2021年2月19日

1. 自己紹介
2. ヘルスケアとIoTとAI
  - 背景、問題点、ヘルスケア特有の事情
  - ヘルスケア適用事例
3. 生体信号と自動車とAI
  - 運転アセスメント
  - 運転能力の変化をとらえる
  - 同乗者の乗り心地
4. 生体信号とゲームとAI
  - 人狼ゲームAIのために
  - 個人適応のための教師データ管理

## As Wright Logic 代表 畑 雅之 (はた まさゆき)

1965年11月19日生まれ。(昭和40年)

北海道教育大学教育学部

公立はこだて未来大学大学院博士課程後期(松原仁研究室)

NECソフトウェア北海道、(財)北海道科学技術総合振興センター、(財)さっぽろ産業振興財団

トライポッドワークス(株)札幌研究センター、株式会社HiSC、東北大学大学院情報科学研究科 研究員

HiSC NORDIC ApS (デンマーク・オールボー市)

現在、As Wright Logic 代表

— UNIXシステム、X-WINDOW、Linux、 $\mu$ i-Tron、JavaVM、MircoDB

ロボットオーサリング、リアルタイムシステム、生体信号センシング、車載生体信号処理、

医療情報解析、ヘルスケアAI、加齢ユーザーインターフェース、糖尿病治療診断システム (AI)、

デジタル治験臨床支援システム

\*PoC前の開発計画や研究計画立案と実施を得意とする



# ヘルスケアとIoTとAI

医療機器データを機械学習で扱う

## IoT はデータ発生源・・・RAWデータは宝の山かゴミの山か？

### ◎増え続けるデータリソース

- センサーの多様化、コモディティ化
- 通信高速化で時間当たりの記録データの増大
- 連続記録、平行記録、補完記録
- 二次生成データ

### ◎大きなデータ

- 静止画、動画、高品質コンテンツの増大
- サイズ、分解能、記録時間

### ◎自動/手動収集データ

- WEBクローリング
- コンバート&ロード
- API連携、サービス連携

### ◎データ収容コスト低下

- 大容量HDD、低価格SSD
- 廉価なクラウドストレージ

### ◎コホートのアプローチの増大

- 集団比較、長期記録観察
- 全記録的プロジェクトの増加

### ◎データ益化/保護

- データ持った者勝ち
- リポジトリサービス
- 記録保管義務
- 学習済みデータ販売
- データクリーニング

## どこにどのように使えば良いか検討もつかない

### ◎データから発想しない

- 価値のあるサービスや機能のありかたや要求から考える
- 必要な情報・データを創出することも検討する

### 【例】自動車の乗り心地を判定したい

**乗り心地の基準は？**

- 車の挙動
- 人の感想
- 人の様子

**いつ判定したい？**

- 作ってるとき
- 商品レポート
- 運転中
- 運転後



\*といっても簡単には見つからない

#### 車からの情報

- OBDIIから
- ドラレコから
- スマホから

ex)

GPS, G, Speed, Temp, Hum  
Movie, Image, Sound,  
Bright,

#### 人からの情報

- ビデオ/音声（乗車時の様子）
- 主観判定（ボタンを押す、事後アンケート）
- 生体信号



そもそもを考えてみる

## • AIを使って価値が増すとは？

- 白黒が付く（判別、鑑別、分類、格付け）
- すぐわかる（リアルタイム、オンライン）
- 推論、推測をしてくれる
- 教えなくて良い（最低限のタグ付け、アノテーション）
- 放っておける（自動、安定、管理負荷が極端に小さい）
- いらぬ物が増える（ストレージがいらぬ、人がいらぬ・・・）
- 人間の限界を超える（稼働時間、注意力、消耗、発見・・・）
- 「人なし」対応（人間が居ない対応）
- AIのせいでできる

# データだけがデータでは無い？

2021年2月19日

AIで付加価値をつけるためには、タグ・アノテーション・時刻・リンクが大事！

- IoTノードから発生するデータ
  - センサー記録値
  - 撮影動画像（音声も）
- 同時に回収したいデータ
  - タイムスタンプ
  - IoTノード情報
  - 環境情報（天候、地勢、イベント）
  - 変調情報（劣化/誤設定/調整）



## オンライン機械学習

### ◎概要

- 逐次与えられるデータを順次学習させて行く方法
- 全てのデータが整う必要は無い
- パーセプトロンが先祖

### ◎特徴

- 学習速度が高速
- 学習後元データを破棄できる
- 学習結果を直ぐに適用できる
- システム化が容易、規模が小さい

### ◎短所

- 学習する順序に依存する
- ノイズによる過学習が起こりやすい
- 元データを捨てる则再検証が難しい

## バッチ機械学習

### ◎概要

- 与えられたデータをまとめて学習させて行く方法
- 全てのデータが整う必要がある
- 様々な学習理論が提案されている

### ◎特徴

- 学習速度がデータ量と適用理論に依存
- 学習後元データを破棄できない
- 学習結果は全ての処理が終わってから
- システム規模が大きく処理負荷が高い
- 商用サービスが整ってきた

### ◎短所

- データの管理が大変
- 規模が大きい場合コストも増大
- 選択できる学習方法の適用見極めが難しい
- 結果の確認までに時間がかかる

## AIにデータを投入前に検討すべき点

### ◎生活記録は厳密な統制が難しい

- 統制するとQoLが著しく低下
- 生活管理や治療への積極さが失われていく
- 生活上の「ノイズ」が含まれている

### ◎QoLが低下しない程度の品質データ

- AIを使って精度を向上と用途拡大
- データフラグメントの補完
- スマホの活用
- SNSの利用
- ラフなデータ取得を許容するデータ仕様

### ◎不連続を許容できるか

- 連続性が必須かどうか
- フラグメントを補完する技術は
- データの発生状況をもとにした近似と補完

### ◎データの管理

- RAWデータのプロテクション（非共同利用）
- 学習データの世代管理
- 特定データの不活性化機能
- オンライン機械学習時の個人情報管理
- 新知見に対する学習の個別更新

## 診断後の個人のデータを学習でとらえる！

### ◎バッチ機械学習からオンライン型に移行

- 診断後は、見るべきは自身のデータ
- 症状の変化に気付く、改善に気付く

### ◎オンライン機械学習で変化に敏感になる

- その日の評価はオンラインで！
- 個人に閉じた学習情報の蓄積 → 治癒経過、変化の観察（抽出）
- 治癒過程の他人との類似性 → 治療方針、治療計画

## 糖尿病の可能性を評価する機械学習システムの応用例

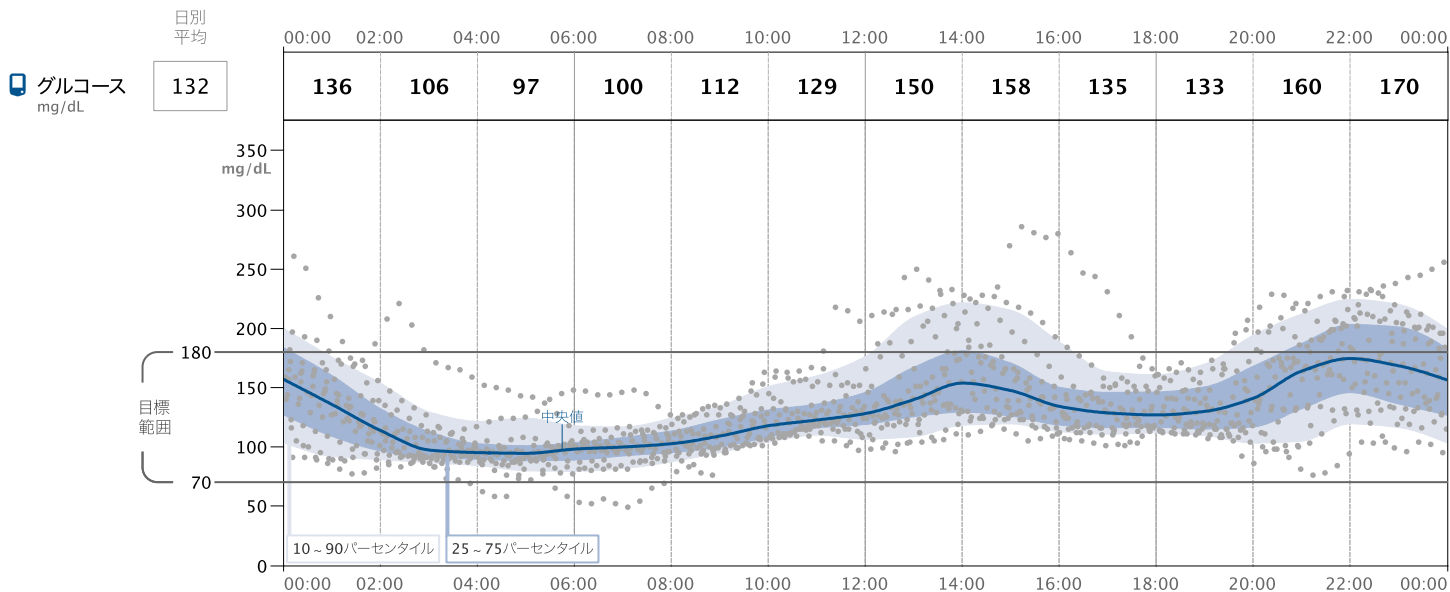
- ◎普及が進む連続式血糖測定器のデータで、糖尿病かどうか評価できないか
  - 訓練データは、14日分を1セット（一人分）× 人数
- ◎15分間隔24時間分のデータで、どの程度の判定ができるのか
- ◎教師データは、糖尿病患者の判定付きデータを使用する
  - 診断指標による判定済み（HbA1c、随時血糖、医師所見）
- ◎健常者の判定付きデータも使用する
  - 健康診断で糖尿病診断指標に異常の無いことを確認
- ◎それぞれ判定済みデータも日別の変化は多様、ルールベースでの判定は難しい
  - 食事内容や運動など、個人の努力による記録情報は断片化しやすく学習に不向き（気まぐれ）
- ◎血糖変化と変化の時間帯のみで、どこまで副次的な知見が得られるのか
  - 将来は、写真の時刻で内容を認識後アノテーション化

15分間隔 / 24時間 / 14日間 の連続した血糖値データを元に  
機械学習によって健常者と糖尿病が疑われるものを判別できるか？

連続式血糖測定器



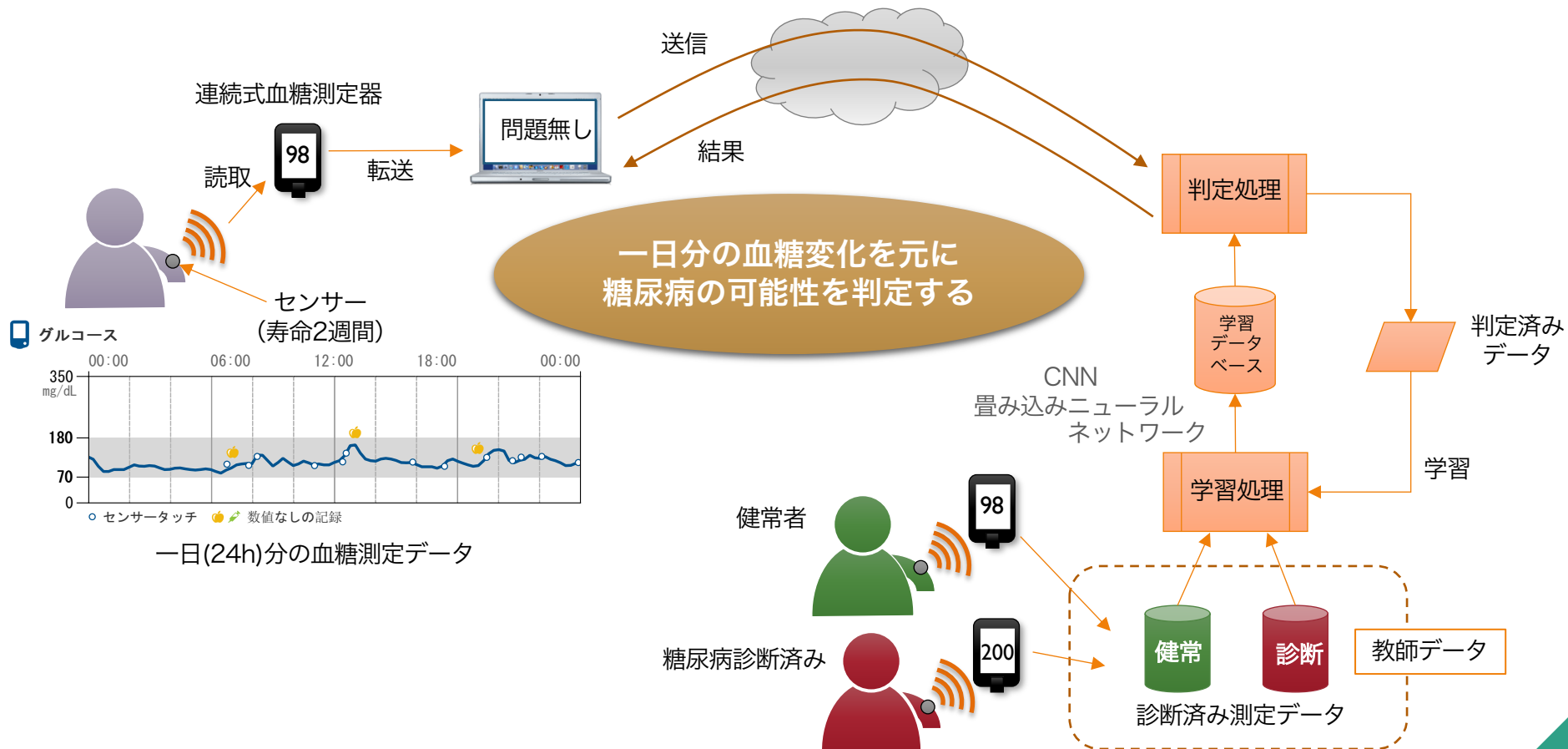
Abbott社製 FreeStyle Libre



参考：Abbott社製 FreeStyle Libre PDFレポートより

# 血糖値情報の機械学習による分析概要

2021年2月19日



# 実験設定

2021年2月19日

血糖値データから健常者か糖尿病罹患/予備群かを判別する  
畳み込みニューラルネットワークの構成を行なう

血糖値データ：14日間分の自動計測データ  
(1日：96データ→1,440データに線形補間)  
※自動計測：15分に1回 体表センサーが記録

## 【注意】

本実験は、医師の診断を代替するために  
行った実験ではありません。

<実験参加者データ：6人分>

学習データを、健常者と糖尿病罹患/予備群の2種類に分類  
健常者

A (50代♂) B (30代♂) C (40代♀) D (40代♀)

糖尿病罹患・予備群

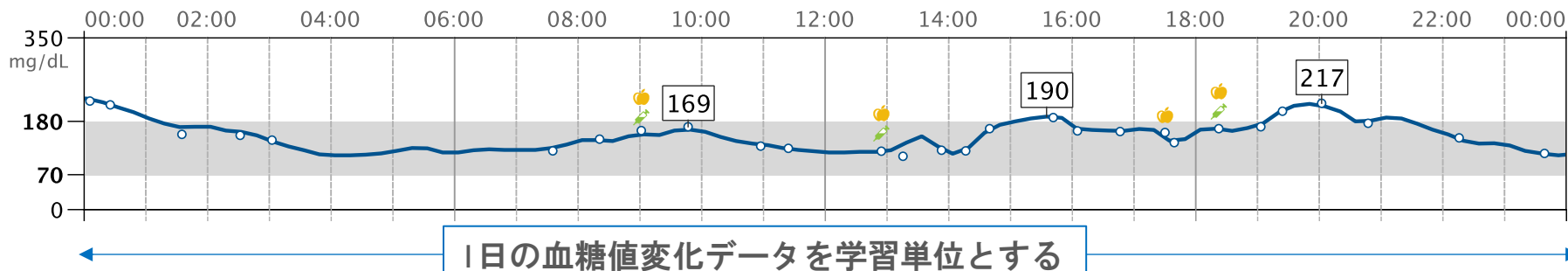
E (40代♂) 、 F (50代♂)

学習データ：A、D、E (一部)、F (一部)

検証データ：B、C、E (一部)

学習と検証を、合計10回実施

比較用に多層NNのみによる学習器を利用した学習・検証を実施



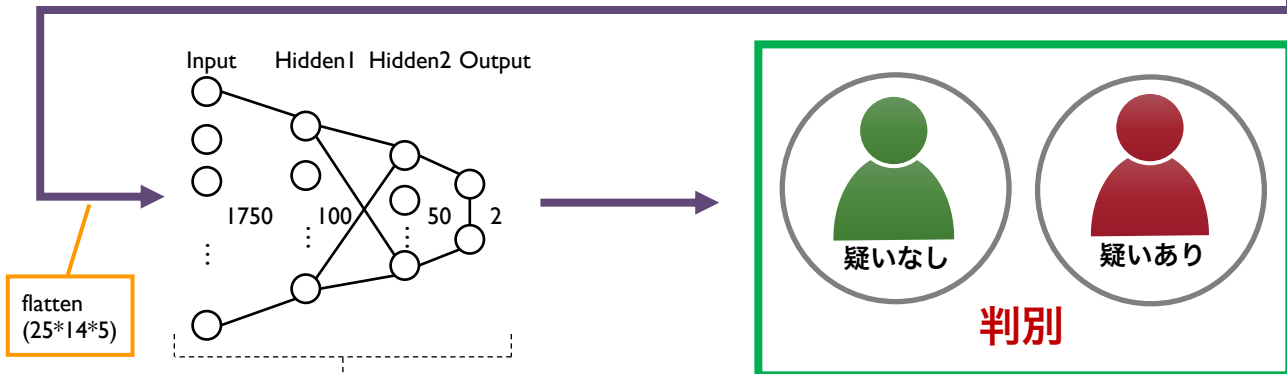
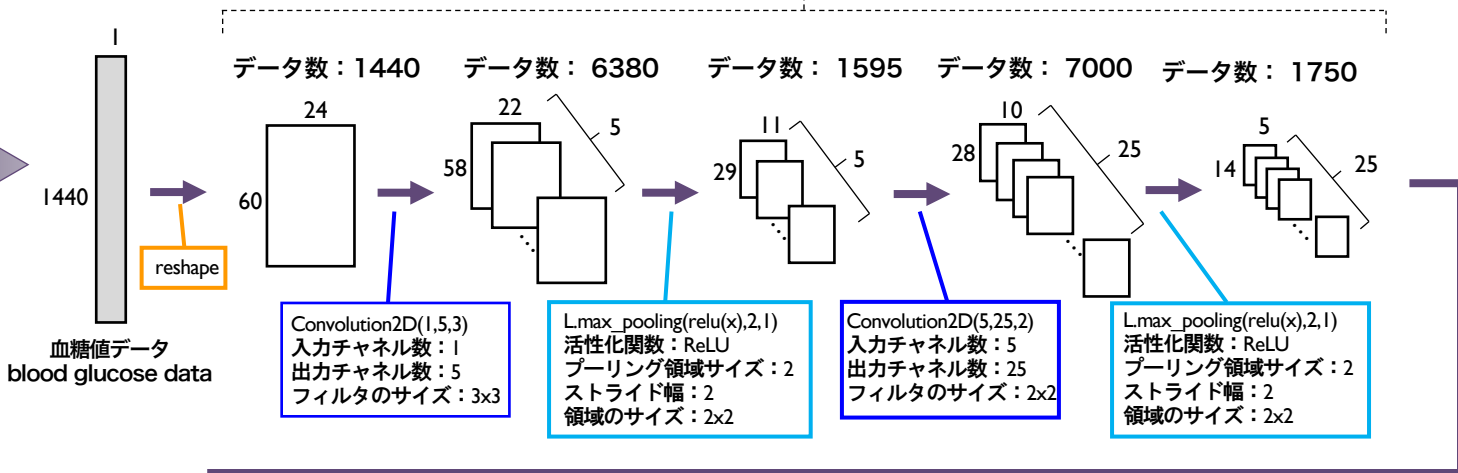


# 畳み込みニューラルネットワークを用いた 血糖値データ判別

2021年2月19日

畳み込み部

1日分の血糖情報  
(15分間隔)



多層パーセプトロン © As Wright Logic

# 学習設定 (CNN、多層NN)

2021年2月19日

## CNN

項目	内容	
データ数	トレーニング	192 うち健常者: 63 糖尿病患者・予備群: 129
	テスト	35 うち健常者: 26 糖尿病患者・予備群: 9
素子数	畳み込み1	1440 (24x60)
	プーリング1	6380 (22x58x5)
	畳み込み2	1595 (11x29x5)
	プーリング2	7000 (10x28x25)
	多層NN (入力-中間1-中間2-出力)	1750-100-50-2
活性化関数	畳み込み1	恒等関数
	プーリング1	ReLU
	畳み込み2	恒等関数
	プーリング2	ReLU
	多層NN入力層	恒等関数
	多層NN中間層1	ReLU
	多層NN中間層2	ReLU
	多層NN出力層	ソフトマックス関数
学習則	Adam ( $\alpha=0.001$ 、 $\beta_1=0.9$ 、 $\beta_2=0.999$ 、 $\epsilon=1e-08$ )	
エポック数	100	
打ち切り誤差	E=0.01	

項目	内容	
畳み込み1	チャンネル数 (入力-出力)	1-5
	フィルタサイズ	3x3
プーリング1	ウインドウサイズ	2x2
	ストライド幅	2
畳み込み2	チャンネル数 (入力-出力)	1-5
	フィルタサイズ	3x3
プーリング2	ウインドウサイズ	2x2
	ストライド幅	2

## 開発環境

項目	内容
OS	Ubuntu 14.10
言語	Python 2.7.8
モジュール	Chainer 3.2.0 scikit-learn 0.19.1 numpy==1.14.0 scipy==0.14.0 pandas==0.22.0

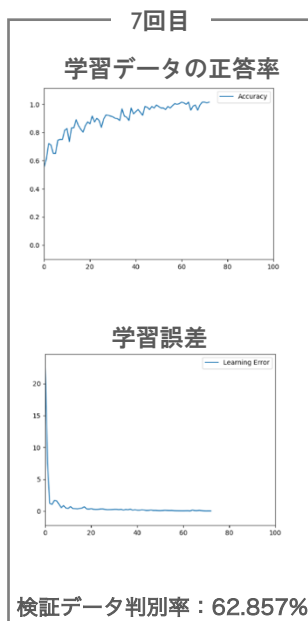
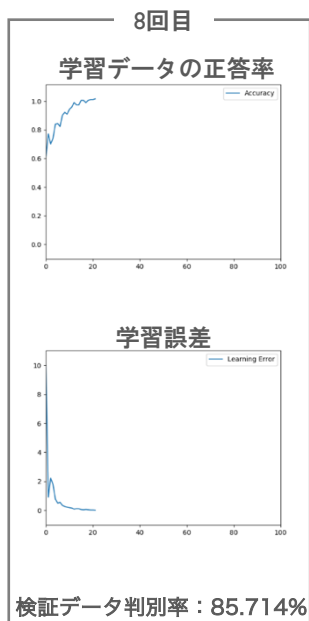
## 多層NN (比較用)

項目	内容	
素子数	入力層	1440
	中間層1	100
	中間層2	100
	出力層	2
活性化関数	多層NN入力層	恒等関数
	多層NN中間層1	ReLU
	多層NN中間層2	ReLU
	多層NN出力層	ソフトマックス関数
学習則	Adam ( $\alpha=0.001$ 、 $\beta_1=0.9$ 、 $\beta_2=0.999$ 、 $\epsilon=1e-08$ )	
エポック数	100	
打ち切り誤差	E=0.01	

# 実験結果

2021年2月19日

判別率	1回目	2回目	3回目	4回目	5回目	6回目	7回目	8回目	9回目	10回目	平均
CNN学習データ	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%
CNN検証データ	82.857%	82.857%	85.714%	82.857%	77.143%	74.286%	62.857%	85.714%	80%	74.286%	78.86%
多層NN学習データ	52.604%	61.979%	51.563%	56.25%	54.688%	64.583%	52.083%	57.292%	66.146%	72.396%	58.958%
多層NN検証データ	45.714%	74.286%	31.429%	37.143%	42.857%	68.571%	47.917%	28.571%	45.714%	65.714%	48.792%



多層NNのみ学習器では、学習がほとんど進んでいない  
(2値なので期待値50%)

CNNでは、検証データの判別率に関わらず、  
学習自体は正常に終了している  
ただし、検証データの判別率が低い場合、  
多いエポック数を要する傾向がある

今回、学習データに対する判別率が100%なので、  
学習データを増やす事で、汎化性能は向上すると考えられる

Adam

$$w_t = w_{t-1} - \alpha \frac{\sqrt{1-\gamma}}{1-\beta} \frac{m_t}{\sqrt{v_t} + \epsilon}$$

$$\begin{cases} m_t &= \beta m_{t-1} + (1-\beta)g_t \\ v_t &= \gamma v_{t-1} + (1-\gamma)g_t^2 \end{cases}$$

初期値

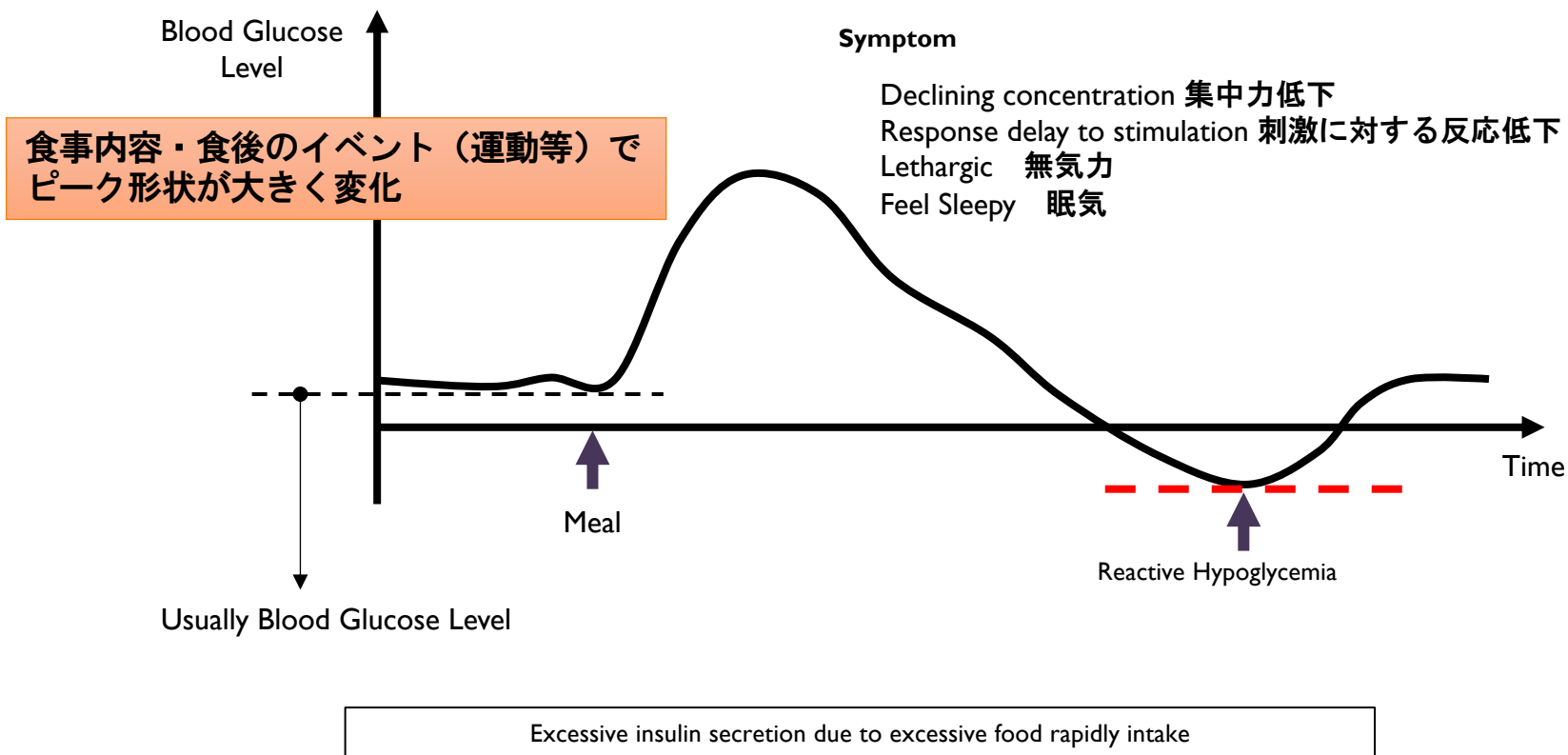
$$\begin{cases} \alpha &= 0.001 \\ \beta &= 0.9 \\ \gamma &= 0.999 \\ \epsilon &= 10^{-8} \end{cases}$$

$g_t$  : 損失関数の勾配

# Reactive Hypoglycemia - 反応性低血糖

2021年2月19日

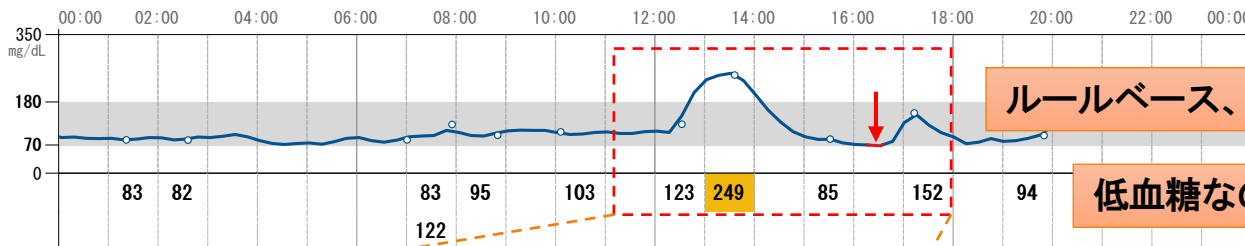
## 閾値やパターンマッチングでは検出が難しい例



# Reactive Hypoglycemia - 反応性低血糖

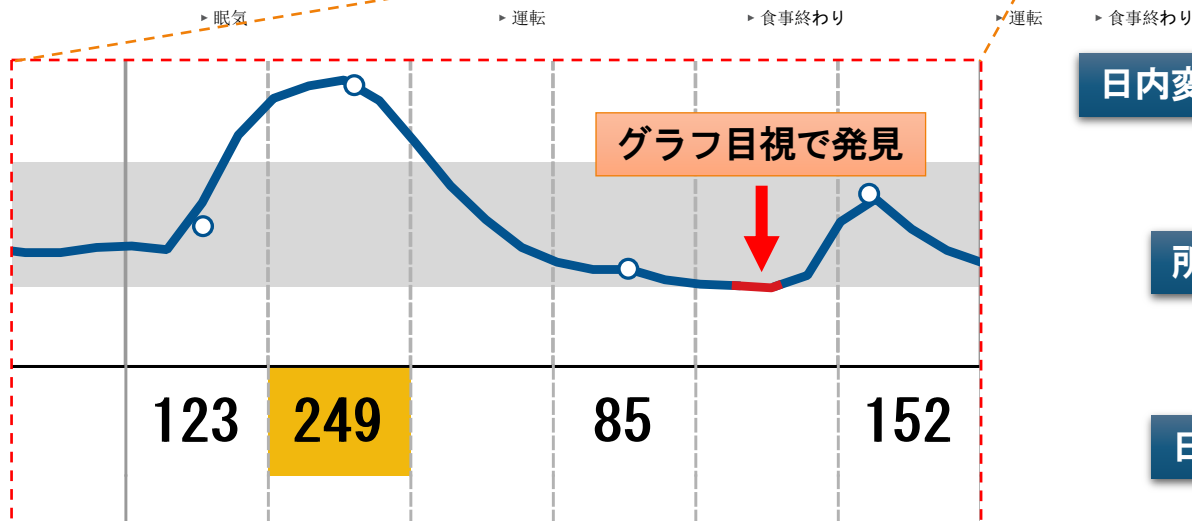
2021年2月19日

健常者にも起こりうる症状をAIでとらえて付加価値増



ルールベース、数理モデルでは難しい

低血糖なのか摂生によるものなのか？



グラフ目視で発見

日内変動全体から機械学習で推定する

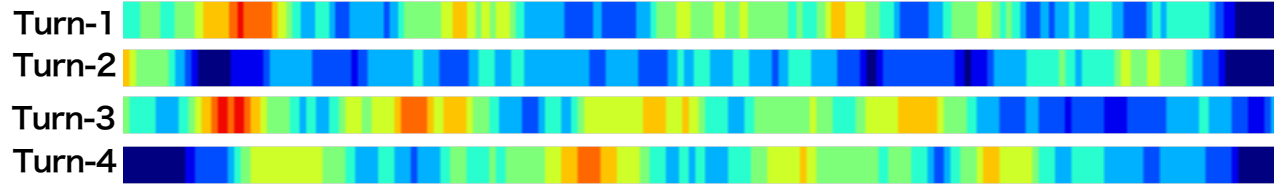
所見付き訓練データを収集する

\*疑似データによる学習

日内に症状があるか推定評価

# 高齢者の屋内行動アセスメント

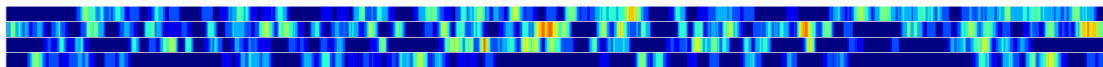
2021年2月19日



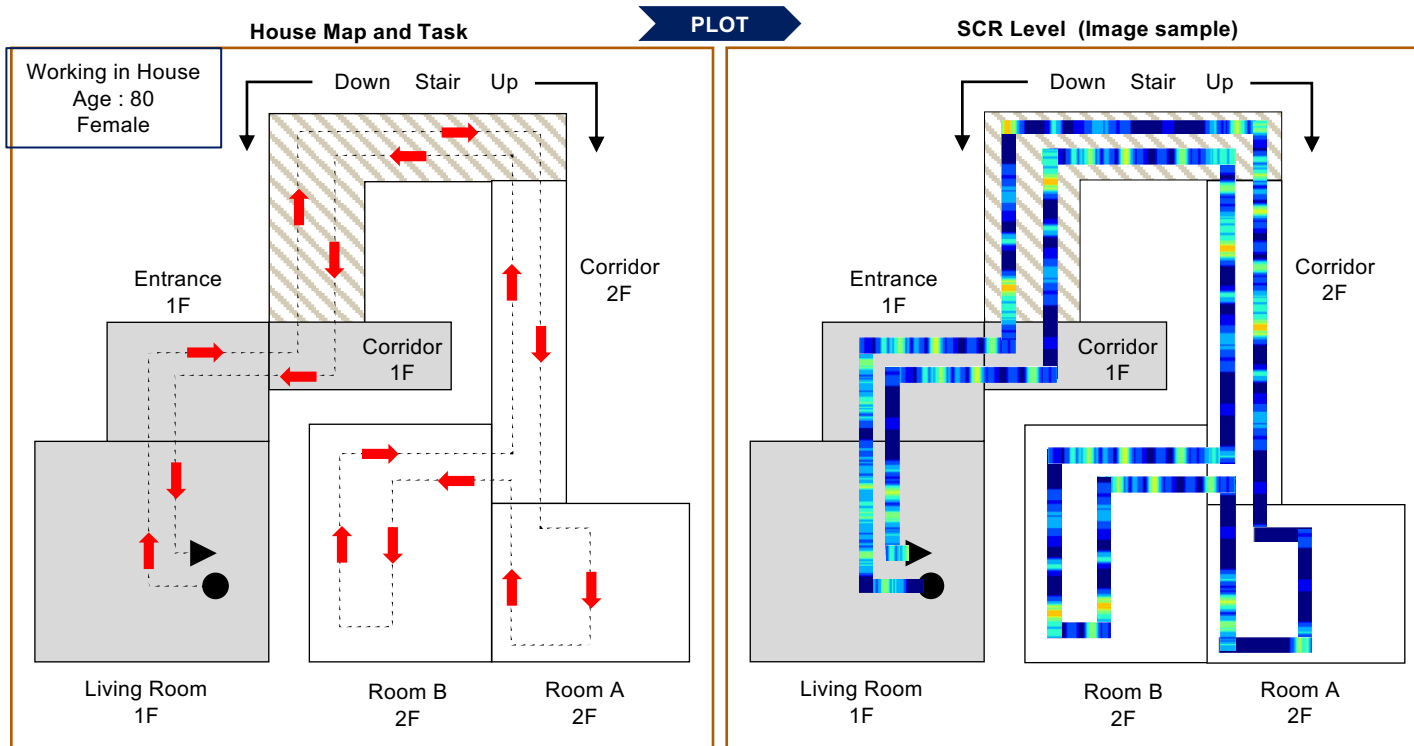
# 生活空間リアクションマッピング

2021年2月19日

SCR Reaction Heat Map TimeLine



No.4



\*プライバシー保護のため実際のデータを一部変更しています。

## 音楽をデータ統制のフレームにした例



時間軸をもとにした学習の適用が容易

TC 00:01:01:05

Turn-1

Turn-2

Turn-3



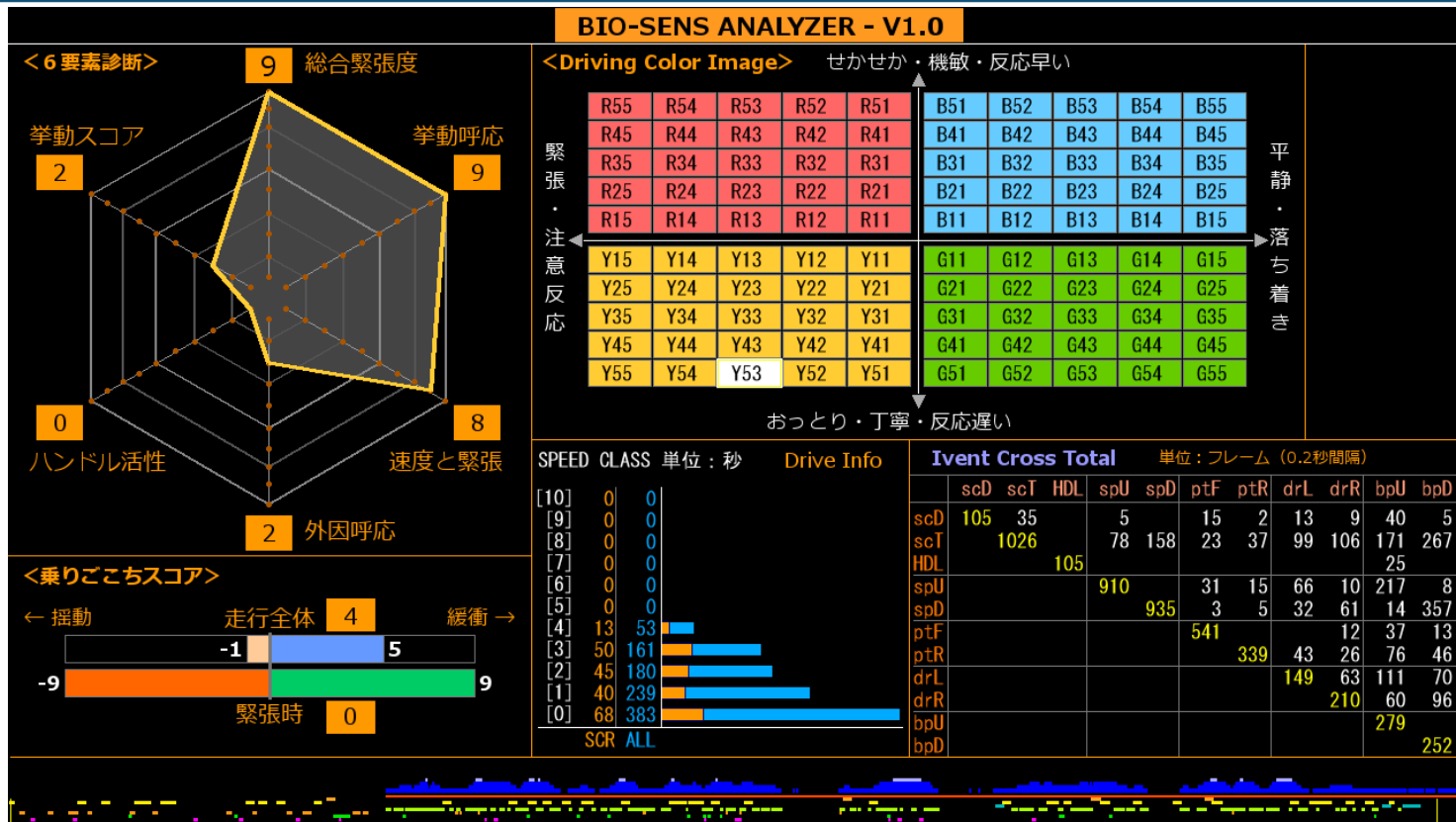
# 生体信号と自動車とAI

生体信号を機械学習で解析

## ドラレコ+SCR(生体信号)から走行中リアクションを解析



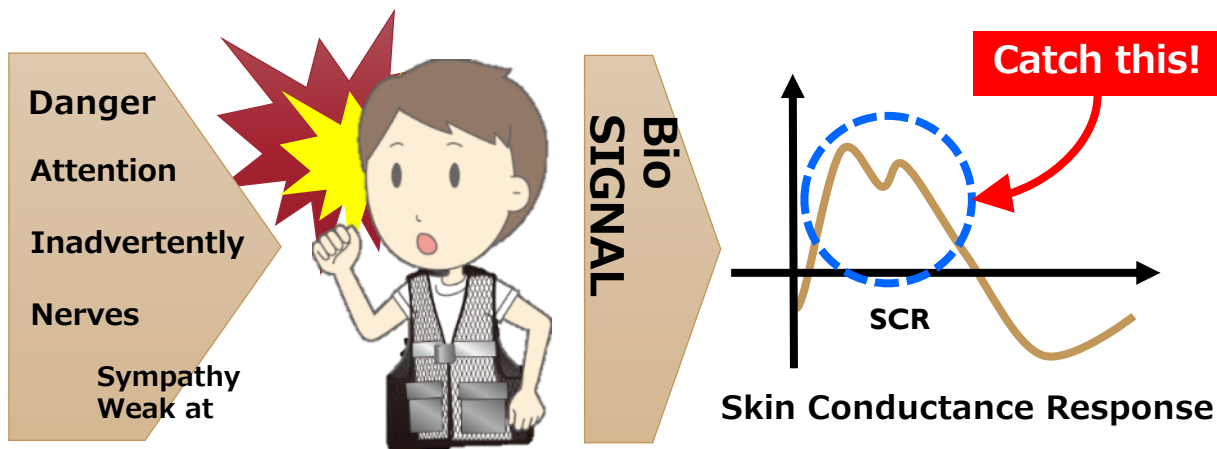
## ドラレコ+SCR(生体信号)から運転傾向をアセスメント - ヒト中心のAI解析



# What is SCR (Skin Conductance Response) ?

2021年2月19日

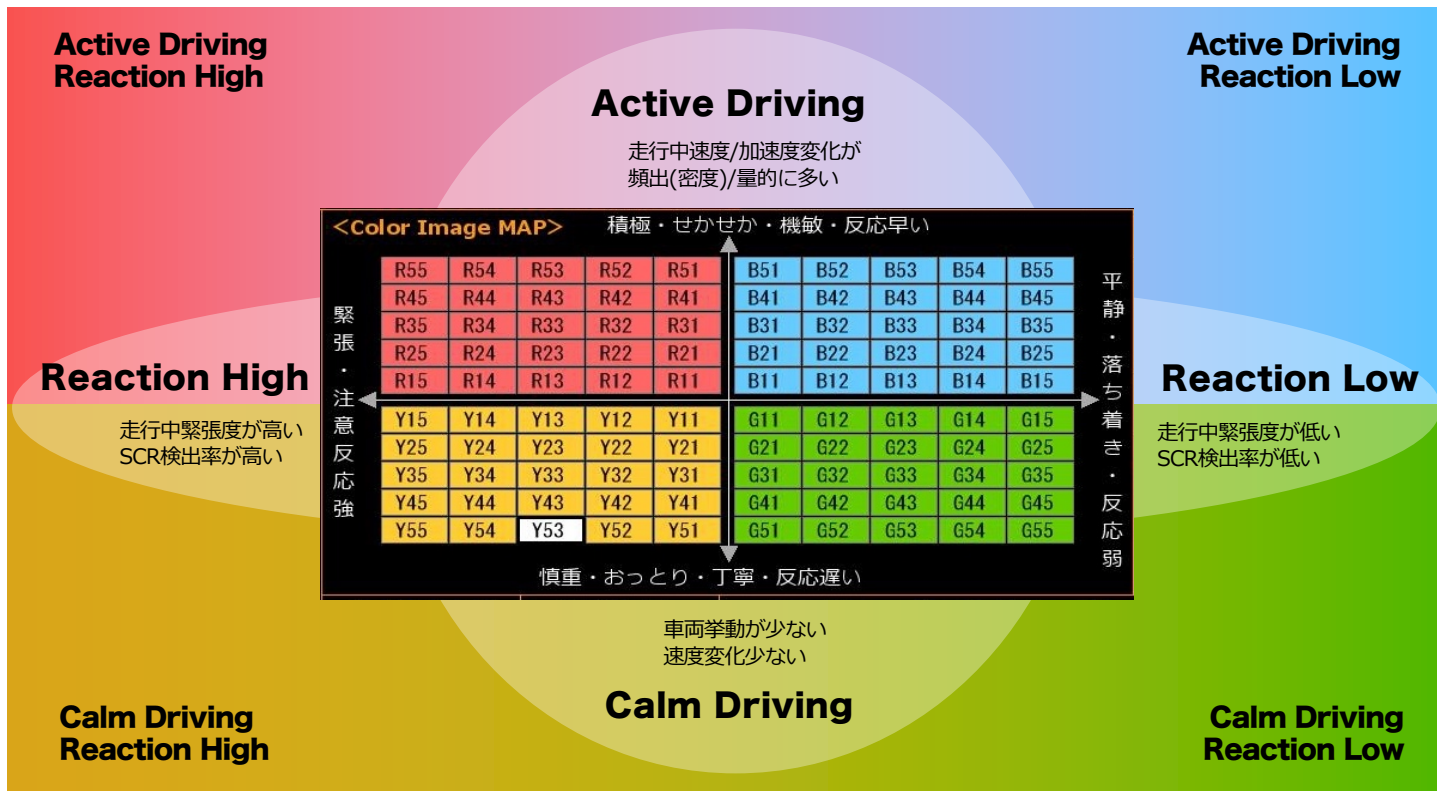
皮膚電気抵抗の変化をとらえ、交感神経の興奮状態を推定する



精神性発汗は交感神経系の緊張や覚醒水準の高さを反映する

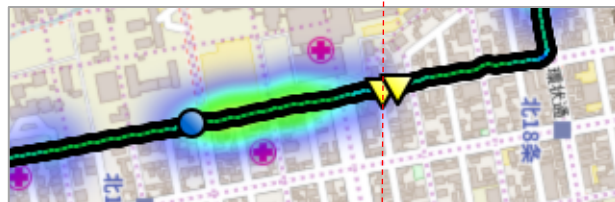
電位は不安定だが波形形状の特徴を機械学習による検出で捉える事が可能

## AIによるアセスメントに必要な評価モデルの構築



# SCR Reaction of Car Event

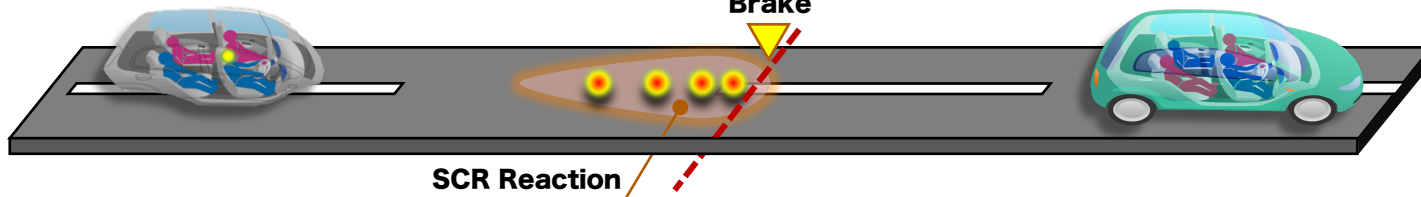
## DRIVER



① Watch >

② Distance

③ Control (Brake)  
**Brake**

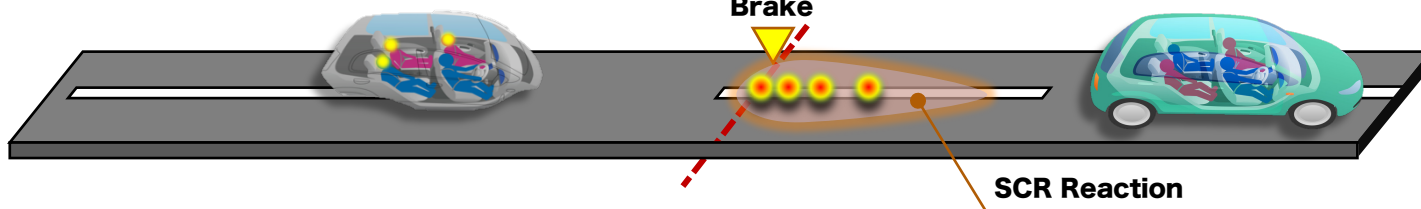


## PASSENGER



① Feel Gravity  
**Brake**

② Reacton



## 運転者の交感神経反応の解析による運転能力評価

詳細情報を俯瞰化した「遅れ」や「ズレ」として表現するところがAI利用のポイント！

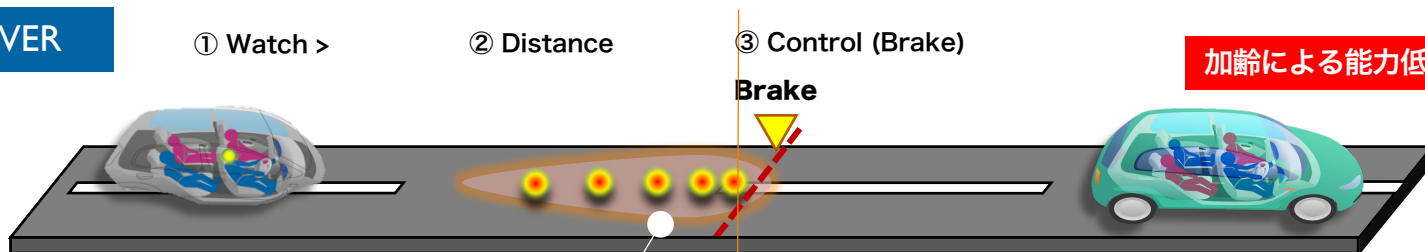
DRIVER

① Watch >

② Distance

③ Control (Brake)

加齢による能力低下の評価



AIによるリアクション評価

SCR Reactions

SCR Reaction

Event Marker  
Through OBDII Interface on Car

Normal reaction  
Preventive

Better

Late reaction

Careful

Late reaction

Worse

Reaction strength

Keep reaction

# Sympathetic reaction on driving

2021年2月19日

運転中同乗者の交感神経の興奮を発汗情報を元に時刻同期して記録・解析

Overlay SCR reaction on the drive movie after testing

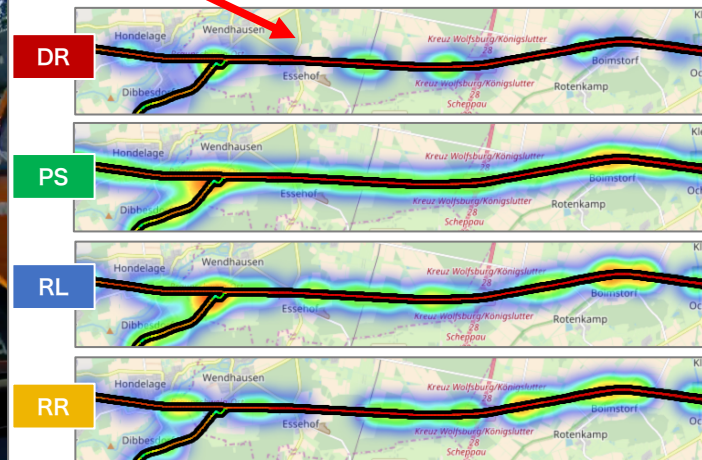
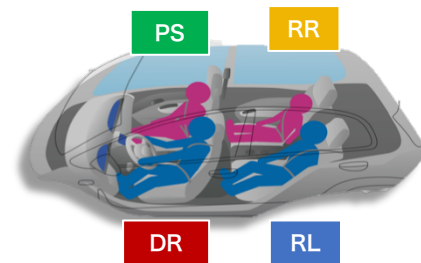
SCR Reaction Signed

機械学習による特徴検出ヒートマップ

DRIVER PASSENGER



REAR LEFT REAR RIGHT



\* Braunschweig Germany  
© As Wright Logic



# Sympathetic reaction on driving

2021年2月19日

走行中の交感神経興奮を捉える

REAR RIGHT



REAR LEFT



PASSENGER



DRIVER

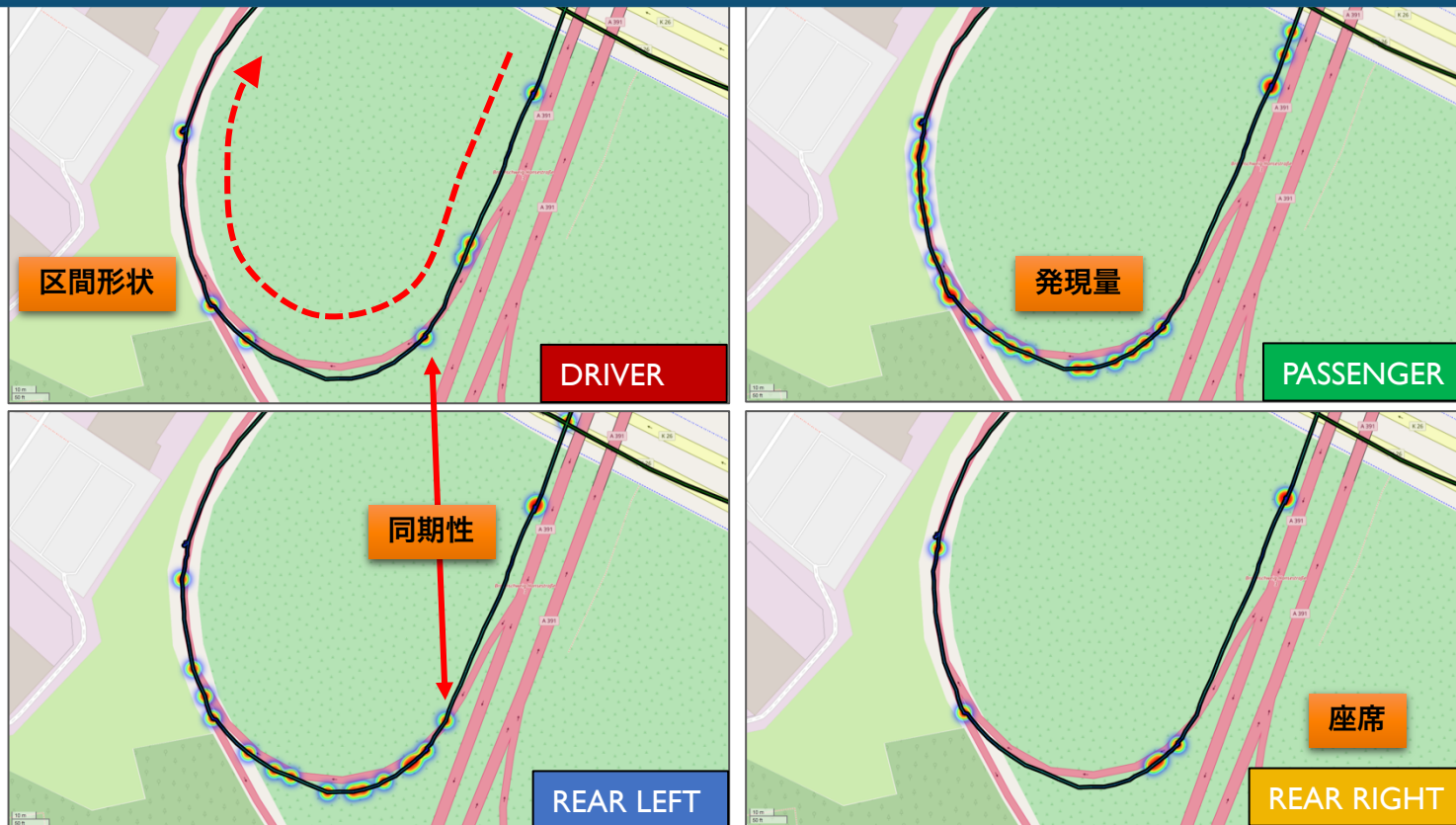


**Digital SCR Logger V4**  
WiFi version



© As Wright Logic

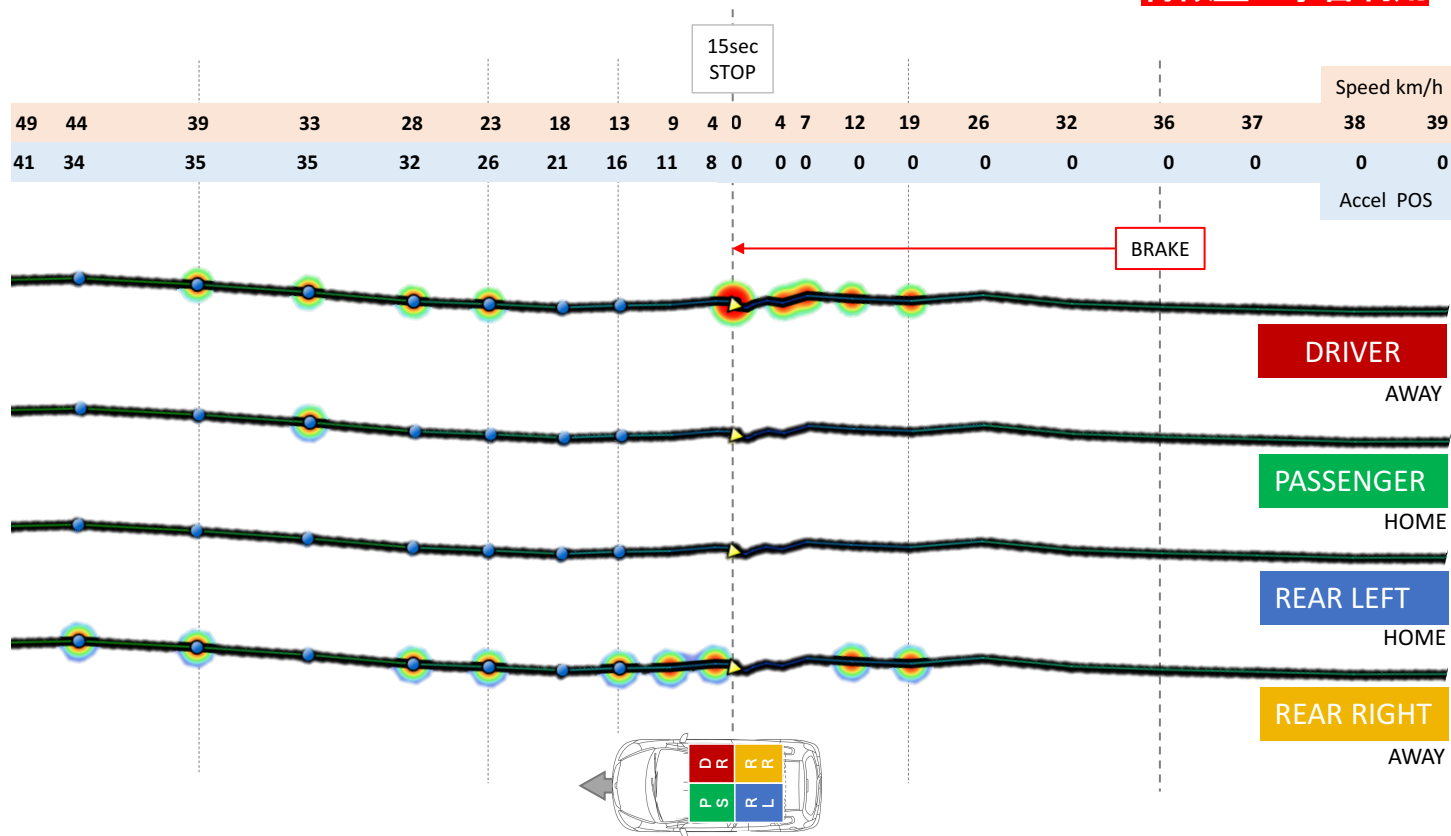
## 緊張度の高い区間のAIによる自動化タグ付けとアノテーション付与



個人  
+  
区間  
+  
区間形状  
+  
時刻  
+  
座席  
+  
車種  
+  
環境  
+  
映像

# イベントリアクションの同期解析 ①

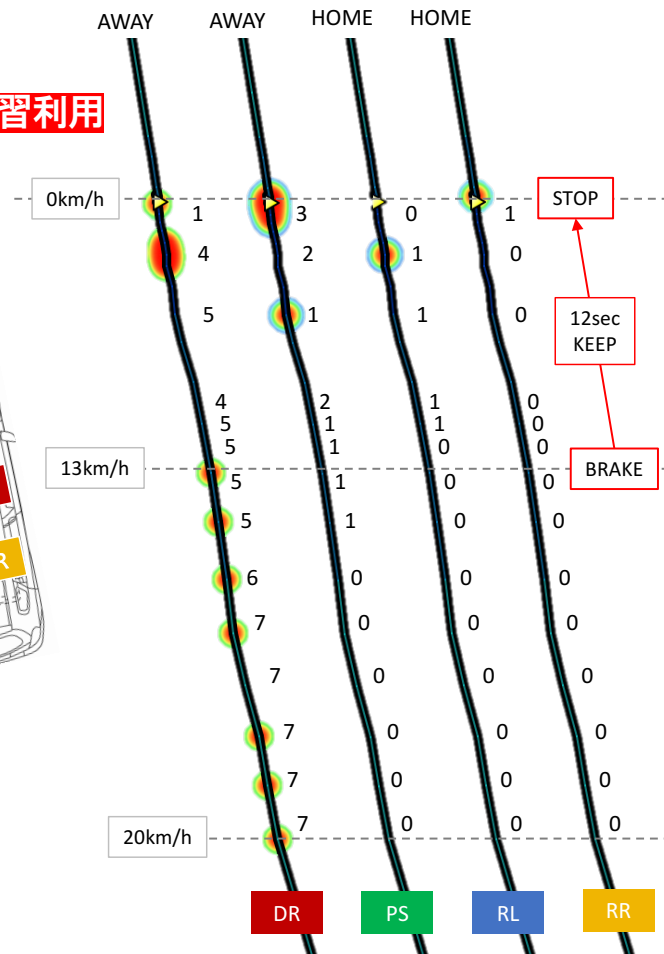
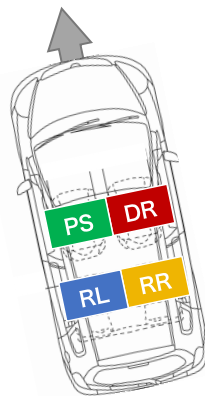
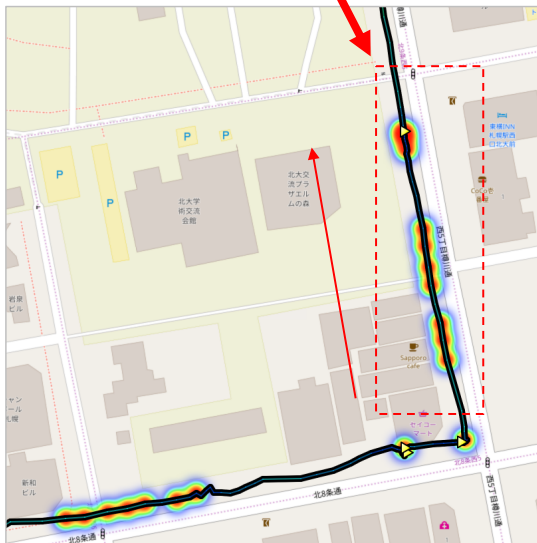
## 特徴量の学習利用



# イベントリアクションの同期解析 ②

## 特徴量の学習利用

## 緊張度の高い走路の学習検出



# 生体信号とゲームとAI

エンターテイメントへの応用

## 機械学習にストレス部位のマークアップ

公立はこだて未来大学/北海道大学との共同研究



汝は人狼なりや？（なんじはじんろうなりや、Are You a Werewolf?）

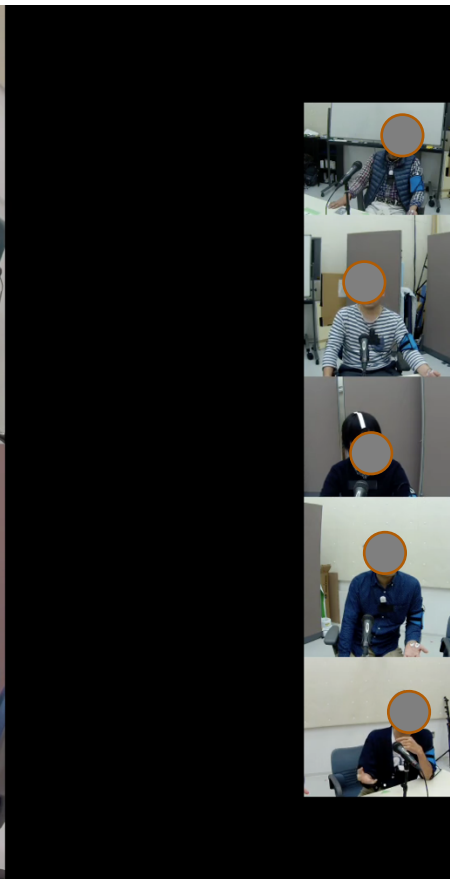
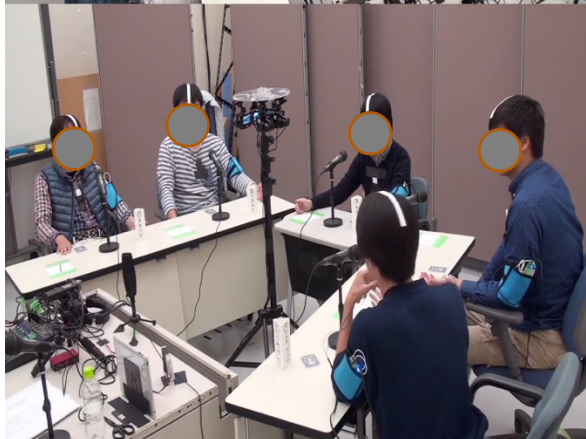
RED:Positive

# Werewolf GAME Reaction

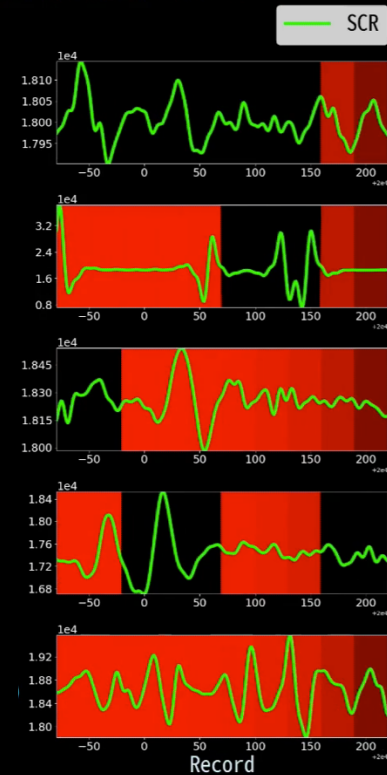
2021年2月19日

Are You a Werewolf?

機械学習による検出部位 **RED:Positive**



2016/11/18 13:42:03



# AsWright Logic