

# AIによる高ストレス者の早期抽出

## —健康診断×ストレスチェックによる予測—

高波龍太郎<sup>1</sup>、西田圭伸<sup>1</sup>、藤井友里<sup>1</sup>、小菅泰雄<sup>1</sup>、  
坂井悠里<sup>1</sup>、清水翼<sup>1</sup>、青柳早苗<sup>1</sup>、大木拓野<sup>1</sup>、  
松井信智<sup>1</sup>、吉田重人<sup>1</sup>

1. IQVIAソリューションズ ジャパン合同会社



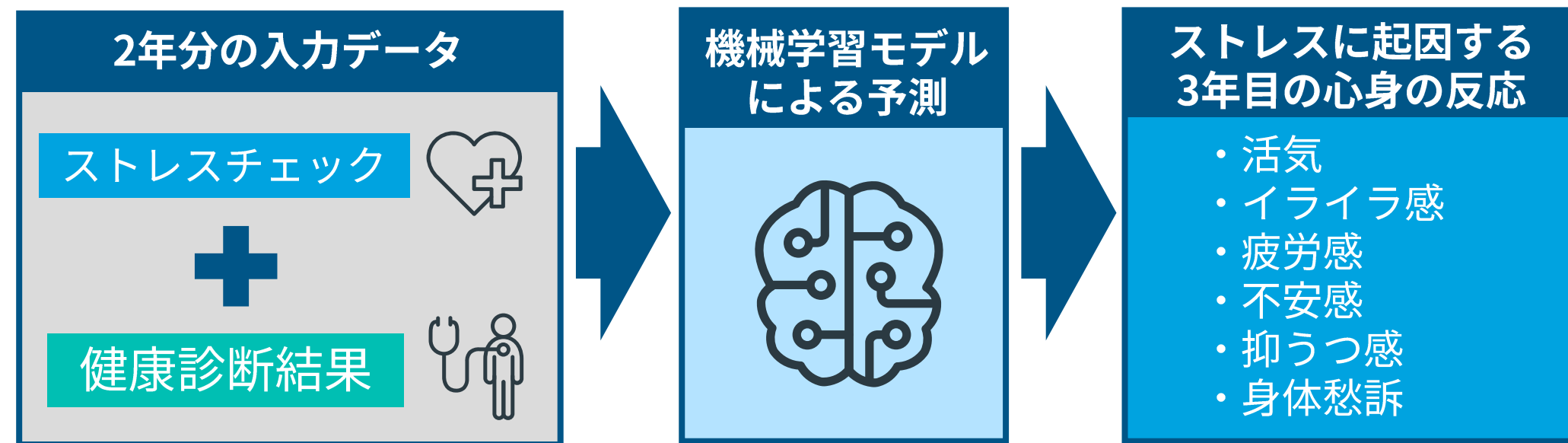
## 背景

職場におけるうつ病や、その他の精神疾患による長期病欠は増加傾向にあり、特に日本ではうつ病関連欠勤の推定コストがブラジル、カナダ、中国、韓国、メキシコ、南アフリカ、米国を含む8か国中で最も高いと報告されている。<sup>1</sup>このため、うつ病リスクの高い労働者を早期に特定・介入する戦略が求められている。日本では1972年以降、労働安全衛生法により常勤従業員への年1回の健康診断が義務化され、さらに2015年からは従業員50人以上の事業所において年1回以上のストレスチェックが義務付けられた。2028年までには従業員50人未満の事業所でも実施が義務化される予定である。簡易職業性ストレス質問票（BJSQ）は、身体的ストレス要因、ストレス反応、職場の社会的支援を評価する57項目からなる、ストレスチェックにおける推奨ツールである。

## 目的

本研究の目的は、企業が保有する年次健康診断データとBJSQデータを統合し、翌年のストレス反応レベルを予測するモデルを構築し、ストレスチェック制度の活用によるメンタルヘルス支援の有用性を検証することである。（図1）

図1：機械学習モデルとデータの関係



## 方法

東京都のA社が提供する複数企業の従業員の年次健康診断データと、BJSQへの回答が突合されたデータベースを用いた。対象は2016年～2023年に連続3年間調査された、18～86歳の従業員8,071人である。予測因子は過去2年間の健康診断データ（年齢、性別、身体計測、血液・尿検査等）とBJSQ回答（57項目）、アウトカムは3年目のストレス反応レベル（活気、イライラ感、疲労感、不安感、抑うつ感、身体愁訴の6項目、1～5点）とした。機械学習モデル(LightGBM および XGBoost) と2年目のアウトカムを予測値として用いるベースモデルとの間で予測精度を比較し、特徴量重要度を評価した。

### 補足説明

**簡易職業性ストレス質問票（BJSQ）**：厚生労働省が推奨するストレスチェックで広く使われる自己記入式質問票（標準57項目）。仕事の負担や裁量などの「要因」、疲労・不安・抑うつ感や身体愁訴などの「反応」、上司・同僚・家族等の「支援（+満足度）」を同時に測り、個人へのフィードバックや集団分析に用いる。

**LightGBM**：勾配ブースティング決定木（GBDT）と呼ばれる機械学習手法の1種である。GBDTは、多数の単純な決定木を逐次的に組み合わせ、予測誤差を段階的に修正しながら精度を高める手法であり、回帰・分類問題に広く用いられている。「非線形な関係」、「多変量データ」、「説明変数間の相互作用」を自動的に捉えられる点が特徴である。LightGBMはGBDTの中でも大規模データに対する計算効率を高めたアルゴリズムである。

**XGBoost**：LightGBMと同様にGBDTの1種であり、より安定性と汎用性に優れる。

**特徴量重要度**：モデルが予測を作るときに各特徴量がどれだけ寄与したかを示す指標。「因果」ではなく「予測上の手がかり」の強さを表す。本研究においては翌年のストレス反応を予測する際に、どの健康診断項目・BJSQ項目が予測に役立ったかを示す。

## 参考文献

- Evans-Lacko S, Knapp M. Global patterns of workplace productivity for people with depression: absenteeism and presenteeism costs across eight diverse countries. *Social Psychiatry and Psychiatric Epidemiology*. 2016;51(11):1525-1537. doi:10.1007/s00127-016-1278-4
- Matsuura Y, Tomooka K, Wada H, et al. The association of long working hours and short sleep duration on mental health among Japanese physicians. *Industrial Health*. 2024;62:306-311. doi:10.2486/indhealth.2023-0174
- Abe T, Okuyama K, Motohiro A, Shiratsuchi D, Isomura M. Association between different types of physical activity and occupational stress in Japanese workers: a cross-sectional study. *Industrial Health*. 2024;62:227-236. doi:10.2486/indhealth.2023-0092

## 結果

最もパフォーマンスが高かった機械学習モデルはLightGBMであった。各アウトカムを高ストレス者（スコア2点以下）とそれ以外に二値化してAUCを評価したところ、すべてのアウトカムで0.8以上を示した。アウトカムの6項目について、全体的に機械学習モデルはベースモデルより高いパフォーマンスを示した。平均AccuracyはLightGBMでは約0.52、XGBoostでは0.50であったが、ベースモデルでは0.48であった。（表1）特徴量重要度は前年のアウトカムや、アウトカムの計算に使用される項目の重要度が高い傾向にあったが、健康診断由来の説明変数の中では、運動習慣や睡眠に関する項目の重要度が高かった。（図2）

表1：各モデルの予測結果

アウトカム	AUC		Accuracy		
	LGB	XGB	LGB	XGB	Base
活気	0.83	0.83	0.51	0.49	0.47
イライラ感	0.81	0.80	0.49	0.46	0.44
疲労感	0.83	0.82	0.54	0.53	0.49
不安感	0.80	0.80	0.50	0.49	0.47
抑うつ感	0.85	0.85	0.51	0.51	0.48
身体愁訴	0.88	0.88	0.56	0.55	0.53

### 補足説明

LGB：LightGBM

XGB：XGBoost

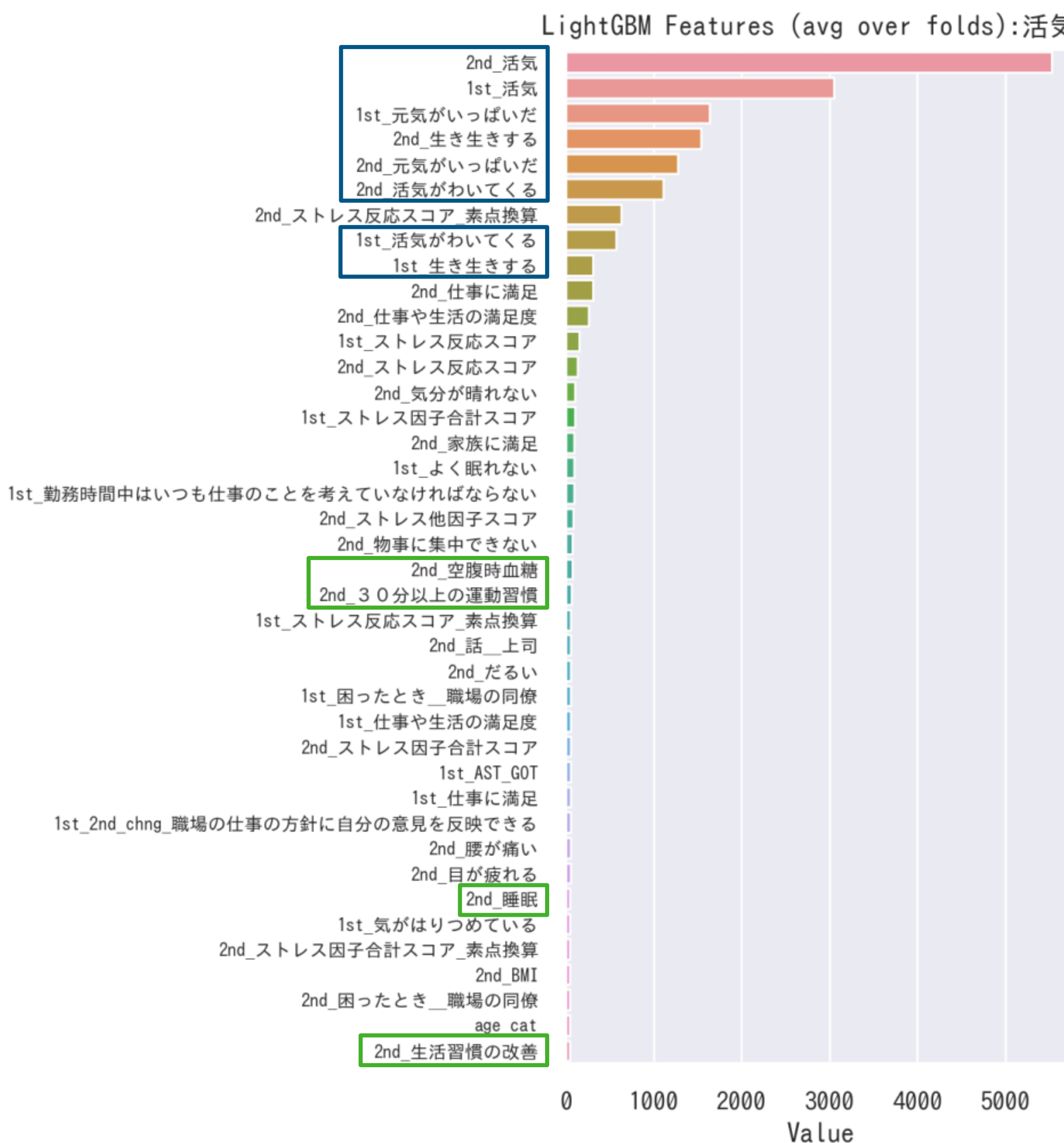
Base：ベースモデル（前年のアウトカムを予測値として用いる）

**AUC**：分類モデルが高ストレス者（陽性）を非高ストレス者（陰性）より高い予測値として識別できるかを示す指標。値は0～1の範囲で、0.5はランダムと同等、1に近いほど識別性能が高い。

**Accuracy**：予測値が実際の値と同じであった割合。0.0（全て不正解）から1.0（全て正解）の範囲。

※各評価指標ごとに、最もパフォーマンスが高いモデルの数値を太字で表示

図2：活気を予測するLightGBMモデルの特徴量重要度



※緑の枠は健康診断の項目から作成された特徴量  
※青の枠は「活気」または「活気」の計算に使用される項目の特徴量  
※上位40の特徴量を抜粋

## 考察

健康診断データとBJSQデータを統合した機械学習モデルは、翌年のストレス反応レベル予測に有用であり、従業員のメンタルヘルス支援に資する可能性が示された。睡眠や運動の影響は先行研究<sup>2,3</sup>とも一致し、予測因子として妥当である。高ストレス者をAUC0.8以上で予測できることから、早期介入や予防措置の推進に役立ち、メンタルヘルス不調に伴う欠勤・休業（アブセンティズム）の抑制に寄与することが期待される。

### COI開示

著者には以下に該当するメンバーが含まれる。「当該企業等の役員、顧問職で報酬額が年間100万円以上」