

# 健診データからの 健康リスク予測技術の開発

サービスイノベーション部 伊藤 拓 落合 桂一†1  
 深澤 佑介†2

近年、従業員などの健康管理を、生産性向上や労働力確保といった経営的な観点で捉える健康経営という概念が重要視されてきている。社員の健康増進が結果として企業の業績アップや価値を高めることにつながるという概念である。その施策として、機械学習を用いた健康予測の実施が進んでいるが、学習モデルがブラックボックスとなっており、なぜそのような予測になるのか説明できないため、ユーザが納得できる解釈を提示することが望まれている。そこでドコモは、健康リスクの予測とその解釈を適切にできるモデル、および予測に基づきリスクを減らすための生活習慣レコメンドを行うサービスを開発した。予測結果と納得感ある解釈により、ユーザが自身の健康と日々の行動との関連を意識し、行動変容を起こして健康増進を図ることを目的としている。健康診断データの活用によりユーザの健康リスクの可視化が可能となり、健康増進に貢献することで健康経営につなげていくことが期待される。

## 1. まえがき

近年、生産性の向上、労働力の確保という経営的な観点から従業員の健康増進を図る、健康経営という概念が重要視されるようになってきた。健康診断や生活習慣に関する問診の結果に機械学習\*1を用い、健康リスク予測を本人へフィードバックすることで、

健康への意識改善を図る取り組みが多く行われている。意識改善を図るためには、予測結果を社員が納得し、生活習慣を改めることが可能な、原因を解釈できる形で出力できる機械学習手法が必要である。しかしながら、機械学習による予測モデルはブラックボックスとなっていることが多く、「なぜ健康リスクがこのように予測されたのか？」という理由を明示する

©2020 NTT DOCOMO, INC.

本誌掲載記事の無断転載を禁じます。

本誌に掲載されている社名、製品およびソフトウェア、サービスなどの名称は、各社の商標または登録商標。

†1 現在、クロステック開発部

†2 現在、総務部

\*1 機械学習：人間が、知覚、経験から知識や判断基準、動作などを獲得していくように、コンピュータにデータから知識や判断基準、動作などを獲得させる技術。

ことが難しかった。既存の解釈手法は、線形回帰<sup>\*2</sup>など解釈が可能な単純なモデルをそのまま予測モデルに当てはめるといったものであった。予測モデルと解釈モデルが一体となっていることで、予測の高精度化と解釈の容易性がトレードオフとなってしまう。この問題を解決するため、XAI (Explainable AI) の概念が登場した。その手法の一例であるLIME (Local Interpretable Model-agnostic Explanations) [1] では、予測モデルと解釈モデルを独立して学習する。

ドコモは、LIMEを利用して社員の健康リスク予測モデルと、予測に基づきリスクを減らすための生活習慣レコメンドを行うサービスを開発した。オリジナルのLIMEに対して、予測の妥当性を高めるための改良を加えたカスタムLIMEを提案し、サービスに導入した。本稿では、提案手法と、本手法を用いた生活習慣レコメンドについて解説する。

## 2. 解釈可能な健康リスク予測および生活習慣レコメンド技術

解釈可能な健康リスク予測のための提案手法では、ある年の健康診断結果と、生活習慣に関する問診結果を入力すると、 $N$ 年後に高血圧・メタボリックシンドローム<sup>\*3</sup> (以下、メタボ) になるリスクと、リスクに寄与している健康診断項目を出力する。以下、リスクを予測する手法と、リスクに寄与している検査項目を計算するための解釈モデルについて解説する。解釈モデルは、オリジナルのLIME、および説明の一貫性を高めるための改良を行ったLIME (以下、一貫性重視LIME) について述べる。

### 2.1 健康リスク予測アルゴリズム

健康リスク予測には、XGBoost (eXtreme Gradient Boosting)<sup>\*4</sup>を用いた、4年間連続して健康診断データが存在するドコモ社員に対して、 $N$ 年分の

検査項目 (11項目) および生活習慣に関する問診 (13項目) を説明変数とし、これを用いて学習させることで、将来の $N+1$ 年、 $N+2$ 年、 $N+3$ 年の検査項目が危険値になっているか否かの目的変数を予測するモデルを、3年分×11検査項目の計33種類構築した。ここで、危険値の定義は厚生労働省の特定保健指導判定値 (2020年3月30日時点) [2] を基準として策定した。危険値の定義を表1に示す。単純に前年度の検査値と同じ値になるとした予測のAUC (Area Under the ROC Curve)<sup>\*5</sup>と比較して、XGBoostによる予測のAUCは7~16%向上した。

### 2.2 LIMEによる健康リスク予測の解釈方法

XGBoostによる予測モデルの解釈を行う、LIMEについて述べる。LIMEの概要を図1に示す。

LIMEでは、予測モデルと解釈モデルを独立して学習させる。予測モデルにはどのようなアルゴリズムを用いてもよい。最適なアルゴリズムにより高精度化が実現できる。しかし、図1のように、例えば行動 $X$ 、 $Y$ から1年後のBMI危険値リスクを予

表1 各検査項目の危険値定義

検査項目	危険値の定義
BMI	25以上
拡張期血圧	85以上
収縮期血圧	130以上
中性脂肪	149以上
GOT	35以上
GPT	35以上
HDLコレステロール	40以下
LDLコレステロール	140以上
$\gamma$ -GTP	50以上
尿酸	7以上
空腹時血糖	100以上

\*2 線形回帰：目的変数が説明変数の係数に対して線形の関係になっている回帰。

\*3 メタボリックシンドローム：内臓脂肪型肥満に高血圧・高血糖・高脂血症のうち2つ以上を合併した状態。

\*4 XGBoost：近年注目されているアンサンブル学習の一種。

\*5 AuC：二値予測問題における評価指標。ROC (Receiver Operating Characteristic) 曲線を描いた際の曲線の面積で、ランダムに予測した場合は0.5、完全に予測が正解した場合は1.0となる。

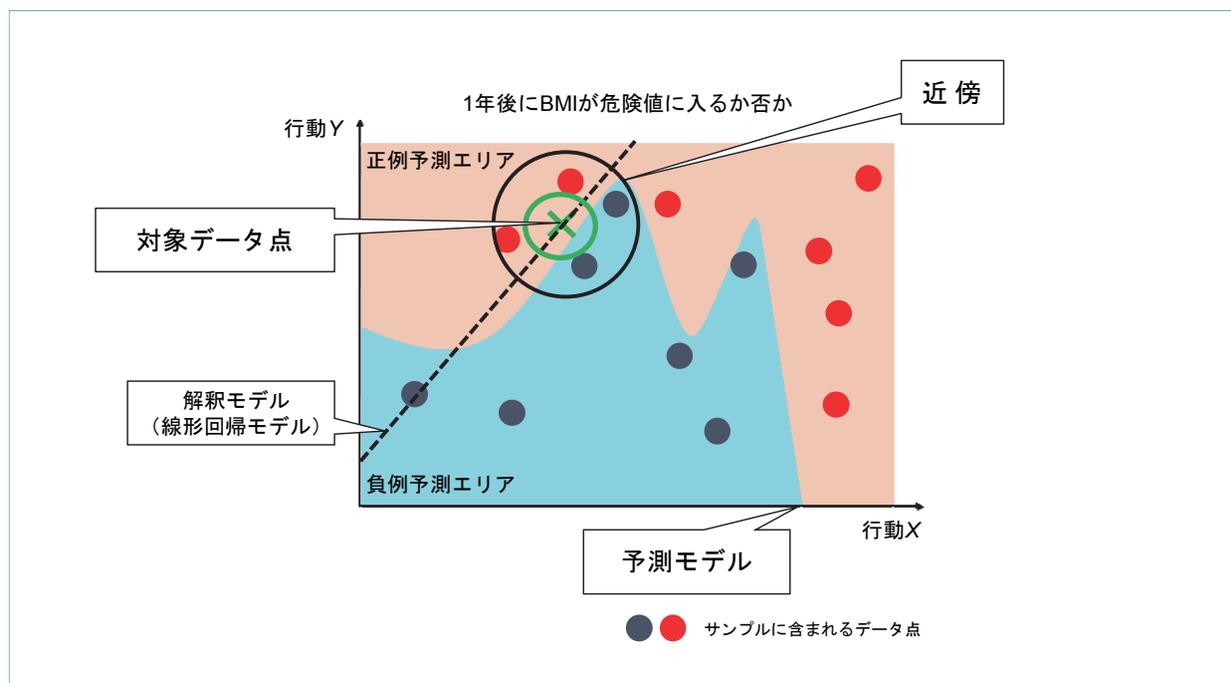


図1 LIMEの概要

測する回帰においてXGBoostなどの非線形モデル<sup>\*6</sup>を用いた場合、そのままでは解釈が困難である。そこで、「何がBMIの値に大きく寄与しているのか？」を説明したいユーザの、あるデータ点について、解釈を行うためのモデルを構築する。解釈モデルでは、説明したいデータ点と、その近傍にある数個のデータ点とを抽出し、それらで予測モデルを近似した線形回帰モデルを構築する。線形回帰モデルであるため、説明変数と偏回帰係数<sup>\*7</sup>の組合せを見ることで、どの説明変数が予測に大きく寄与しているかを読み取ることができる。このように、説明したいデータ点近傍で、予測モデルに近似したモデルにより解釈を行うため、データ点によって、生成される解釈モデルは大きく異なる。つまり、健康リスク予測では、予測対象の社員によって、健康リスクに寄与する生活習慣が異なる結果となることを意味し、個人に合わせた解釈を行うことが可能となる。

### 2.3 説明の一貫性を重視したLIMEのカスタマイズ

一貫性重視LIMEの概要を図2に示す。

#### (1)従来のLIMEの課題

従来のLIMEの場合、説明したいデータ点の近傍で近似モデルを作成するので、解釈モデルは抽出する近傍のデータ点に依存する。そのため、極端に一般的な傾向から外れたデータ点がある場合には、その外れ値によって解釈が歪められてしまう危険性がある。例えば、ある年に突然の疾患により一時的に体重が増加した、または血圧が上昇した、というサンプルが含まれていた場合、生活習慣が良くなかったゆえ疾患になったケースもあれば、生活習慣に関係なく発症してしまうケースもあるため、ランダムに近傍データ点を抽出する従来のLIMEでは、後者のデータ点も含んで解釈モデルを生成してしまう可能性があり、その場合生活習慣を改善させるための正しい解釈モデル生成はできない。

\*6 非線形モデル：目的変数と説明変数の係数が、非線形な関係で表されるモデル。非線形モデルの方が一般的に表現力は強い。

\*7 偏回帰係数：回帰分析において、得られる方程式における説明変数の係数。

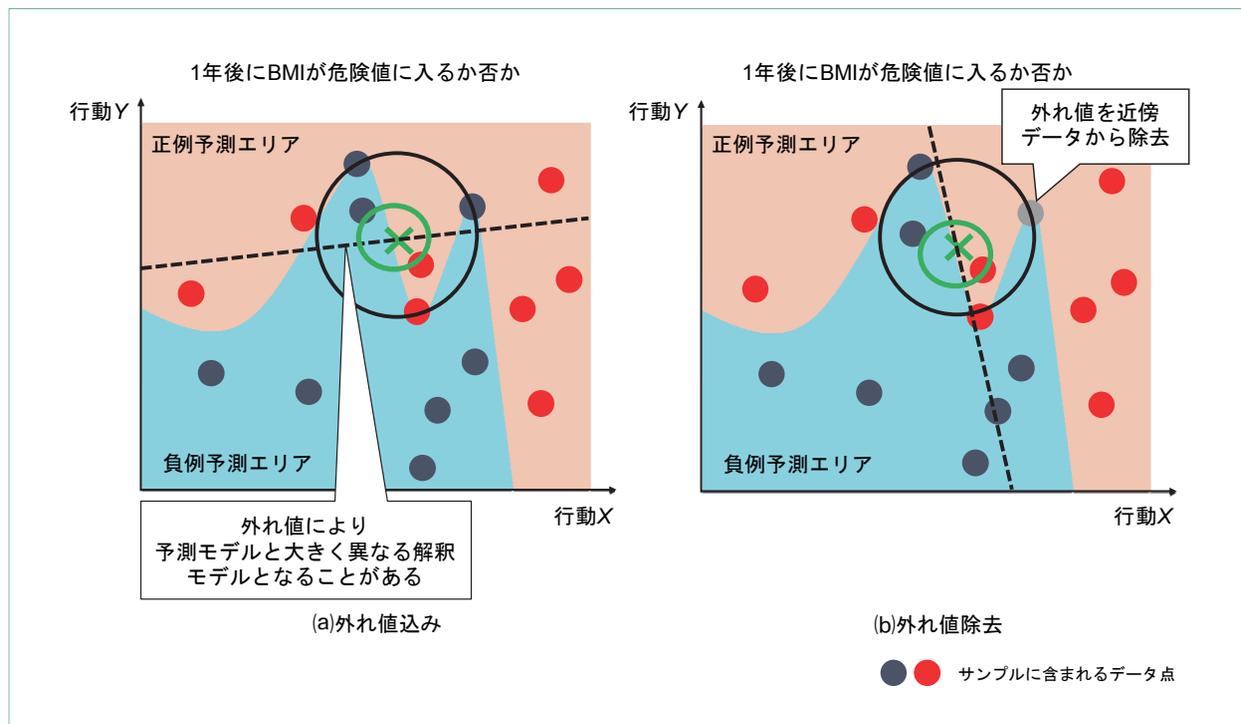


図2 一貫性重視LIMEの概要

(2)一貫性重視LIME

一貫性重視LIMEでは、予測の妥当性を高めるために改良を行った。具体的には、説明したいデータ点に対して近傍データ点を探索する際に、近傍点が外れ値でないかをチェックし、外れ値の場合それを除外する機構を追加する。

チェックにあたっては、まず説明したいデータ点と、近傍データ点候補との間の直線の傾きを計算する（以下、傾きA）。また、説明したいデータ点における予測モデルの微分係数を計算する（以下、微分係数B）。データ点の近傍で説明変数の特性が一致している場合、傾きAと微分係数Bは正負が一致していると考えられる。例えば、飲酒量と1年後のBMI予測値を考える際、説明したいデータ点の近傍では、「飲酒すると1年後のBMIが上昇する」という特徴があった場合、近傍で抽出されたデータ点でも、同じような傾向を示すはずである。しかし外れ値が

ある場合、予測モデルが描く曲面と大きく離れた位置にデータ点が存在する。そのため、傾きAと微分係数Bの正負が一致しないケースが発生する。したがって、この2つの値の正負が一致しないデータ点については、周辺のデータ点と傾向が大きく異なる外れ値であるとみなせるため、解釈モデルを生成する際のサンプルからは除外する。

外れ値を含む解釈モデルと、外れ値を除外した解釈モデルの比較を図2に示す。左図は、外れ値により予測モデルと大きく異なる解釈モデルが生成されているが、右図は外れ値を除去したことにより、予測モデルに近い解釈モデルを生成できている。以上が一貫性重視LIMEのアルゴリズムである。

(3)オリジナルLIMEと一貫性重視LIMEの比較結果

オリジナルLIMEと一貫性重視LIMEのリスク予測結果の違いを表2に示す。この表は、同一のサンプルに対してオリジナルLIMEと一貫性重視LIME

表2 オリジナルLIMEと一貫性重視LIMEによるリスク予測結果の変化

&lt;オリジナルLIME&gt;

1年後のBMIへの影響		2年後のBMIへの影響		3年後のBMIへの影響	
生活習慣	影響スコア	生活習慣	影響スコア	生活習慣	影響スコア
飲酒量	-0.14	身体活動	-0.17	体重増	0.08
飲酒習慣	0.09	早食い	-0.08	早食い	0.03
身体活動	-0.06				
一貫性のない予測					

&lt;一貫性重視LIME&gt;

1年後のBMIへの影響		2年後のBMIへの影響		3年後のBMIへの影響	
生活習慣	影響スコア	生活習慣	影響スコア	生活習慣	影響スコア
就寝前夕食	0.05	就寝前夕食	0.04	睡眠	0.03
飲酒習慣	-0.03	身体活動	-0.04	飲酒量	0.03
		睡眠	0.03		

を適用し、1~3年後それぞれのメタボリスクを予測し、そのリスクへの寄与度が高い生活習慣を比較したものである。オリジナルLIMEでは、2年後と3年後の「早食い」における寄与度が正負逆になっている。通常であれば、早食いの習慣を続けた場合に2年後のBMIは下がるが3年後のBMIが上がる、などとは考え難い。このような結果となったのは、近傍データ点の中に、2年後もしくは3年後どちらかに急激なBMI変化が記録されたサンプルがあったために、その影響を受けてしまったためと考えられる。一方で、一貫性重視LIMEではそのようなデータ点を排除しているため、上記のような直観に反する予測は行われていない。納得感のあるリスク予測結果を提示するためには、このような妥当性は重要であると考えられる。

### 3. 生活習慣レコメンド技術のサービス展開

上記技術を、ドコモのdヘルスケア上のサービス

に展開し、NTTグループの社員が、自身の健康リスクとそのリスクに寄与している生活習慣を確認できるサービスを昨年4月にリリースした。図3にサービス画面を示す。

サービスとしては、NTTの健康保険組合に加入している社員が、「NTT健康ポータルナビ」に登録されている自身の健康診断結果をdヘルスケアに転記する許可を行うと、dヘルスケアにデータが連携され、上記計算を行うライブラリ<sup>\*8</sup>を介して、健康リスクと、そのリスクに寄与している生活習慣を出力するというフローになっている。ライブラリ内のモデルは、ドコモが保有する数年にわたる10,000以上の健診データを基に学習を行った。

## 4. あとがき

本稿では、健康経営の重要性の高まりの中にあって、機械学習を用いた健康リスクの予測と、その解釈を行うためのモデルにおける、困難性と解決するためのアプローチについて解説した。

\*8 ライブラリ：汎用性の高い複数のプログラムを、再利用可能な形でひとまとまりにしたもの。



図3 dヘルスケアにおける実際のサービス画面

実際に社員向けにサービス展開されており、また社外に向けては法人向けの健康経営支援サービスの提供も開始されたことから、今後も技術改良を加えながら、社員の健康増進に貢献するとともに、社外へのドコモの健康経営のアピールの一助としたい。

## 文献

- [1] M. T. Ribeiro, S. Singh and C. Guestrin: "“Why Should I Trust You?”: Explaining the Predictions of Any Classification," KDD'16, 2016.
- [2] 日本予防医学学会: “健康診断結果の見方: 特定健康診断 (メタボリックシンドローム).”  
<https://www.jpml1960.org/exam/exam01/exam15.html>

※本データはHealth Data Bankのデータを活用しており、本人に対して利用目的を明らかにし、同意を頂いた上で取得したものである。