

滋賀大学における
データサイエンスの産官学連携 (9)

人的資本分析における 機械学習の実務応用

岡田 浩一



NTT
ExC Partner

Human Capital
Analytics

近年、機械学習技術の進歩に伴い、人的資本分野でのデータ活用に機械学習を取り入れる動きが注目されている。この分野で研修やコンサルティングを提供しているNTT ExCパートナー (hc-analytics-ml@nttexc.co.jp) では、滋賀大学と連携して、従業員のパフォーマンスや行動の分析等において機械学習の導入を検討し、経営戦略や人事施策への活用に取り組んできた。特に、機械学習モデルの説明機能 (Explainable AI、XAI) によって、どの要因がどの程度影響しているかを示す方法を調査し、効果的な施策設計や対象絞り込みを可能にした。本稿では、人的資本分析の1つである「エンゲージメント分析」を例として、機械学習の実務応用方法とその特徴を解説する (図1)。

エンゲージメントは従業員の仕事に対する熱意や満足度を示し、組織の生産性や離職率とも関わる重要な指標である。この分析では、エンゲージメントを目的変数として、その増減に影響すると考えられる複数の説明変数 (先行指標) を設定する。これらの説明変数は、何らかの施

策によって改善可能であることが特に重視され、アンケート結果や勤務時間、上司評価などから取得される。

次に、これらのデータを用いてLightGBMなどで非線形の機械学習モデルを構築する。非線形モデルは複雑な変数間の関係性を捉えやすく、単純な線形回帰より予測精度を高めることができる。その反面、モデルの内部挙動がブラックボックス化しやすいため、結果の説明が難しくなる。そこで有効なのがLIMEによる局所勾配の算出である。LIMEは複雑なモデルの動きを局所的に線形モデルで近似し、個別の従業員データに対してどの説明変数がどれほど結果に寄与しているかを明らかにする。例えば、ある従業員については「上司からの評価」がエンゲージメントに大きく影響している一方、「残業時間」はあまり寄与していないといった詳細な分析が可能である。

従来の解析手法である重回帰分析では、分析者が従業員を特定のグループに分けて傾向を調査することが多く、観察者の仮説や偏見が結果に影響するリスクがあった。これに対し非線形モデルは、人をグループ分けせずデータ全体の傾向を学習し、その後LIMEなどの説明手段で変数の単独効果や複合効果を確認することで、より客観的かつ網羅的な分析が可能になる。

説明変数の寄与率の評価は単純な全体平均や従業員ごとの個別値だけでなく、説明変数ごとの分位点別や従業員セグメント別で効果を示すことも効果的である (図2)。例えば、年齢の若い20%の層には施策が著しく効果を発揮し、高い層には影響がほとんどないと分かれば、限られたリソースをより効果的に使える。

図1 機械学習の実務応用プロセス

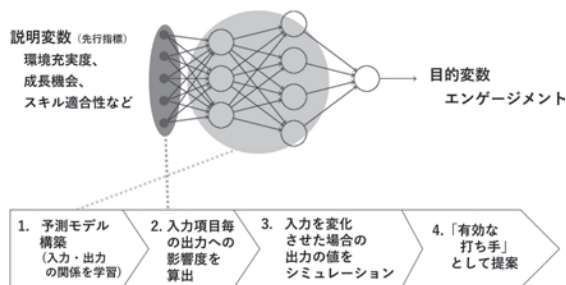
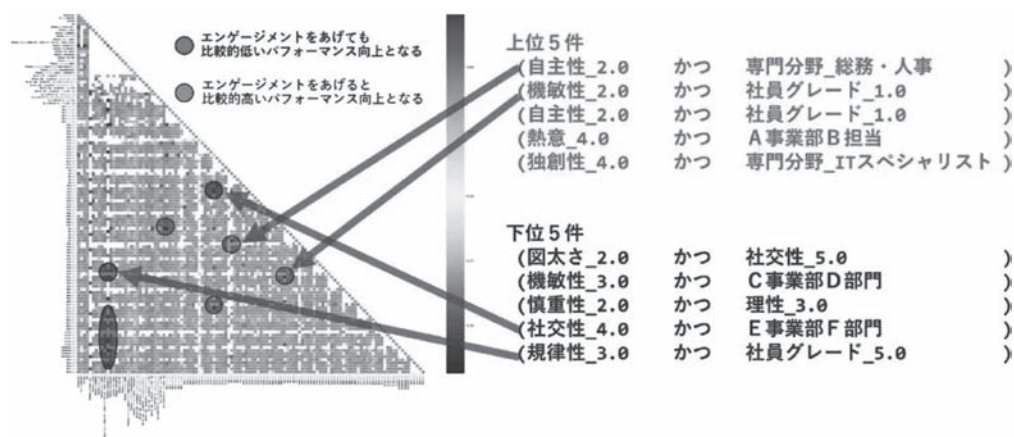


図2 LIMEによる局所勾配の見える化の例



目的変数であるエンゲージメントの改善にむけて、例えば「研修受講の促進」や「業務環境の改善」といった施策が説明変数の値に介入することを狙って実施される。このとき、施策の効果とコストのバランスを考慮して、効果が出やすい施策を特定することが有効である。そのために機械学習モデルを使ったシミュレーションで、各変数の改善が組織や個人の指標にどう影響するかを事前に評価する。これにより、施策の優先順位を定められる。

単純に重回帰分析の係数情報だけで介入を決める場合でも、効果を過大評価するリスクがある。例えば、既に高い値を持つ従業員が多い変数に介入しても、大きな改善効果は期待できない。従って、改善の実現可能性やコストを踏まえて変数の改善値を調整し、その上で介入対象に対して介入後のエンゲージメントの予測を行うことで、効果を評価することが望ましいと言える。また、非線形モデルでは、ある変数の介入効果が他の変数の状態に依存して変化するため、介入後の説明変数を使った再予測によって効果の有無や逆効果がないかを慎重に確認する。

また、説明変数間の連動（相関）にも注意が必要である。例えば、残業時間が減れば職場評

価など他変数も変わることが考えられる。介入変数と連動する変数は、過去データの相関の確認や、現場の知見を元に検討する。連動の強い変数は多重共線性（説明変数間の過度な相関）を発生させ、モデルの安定性や解釈性を損なうため、対象領域の業務知識に照らして片方の変数を除外することを検討する。

さらに交絡因子の存在確認も重要である。交絡因子とは、目的変数と説明変数等に同時に影響しうる第三の要因で、これを考慮しないと予測の誤差が大きくなる。交絡因子が存在すると想定される場合、説明変数の見直し等が必要になる。

こうした分析によって、実務でのデータ活用価値が高まる。総じて、人的資本分析においては、高精度な予測モデルの構築だけでなく、LIME等の局所説明手法によるモデル挙動の可視化やシミュレーションによる施策設計、現場知見との往復検証が有効である。これにより、企業は人的資本の持続的な向上と競争力強化を図ることが期待される。

(おかだ こういち
株式会社NTT ExC パートナーマーケティング部
HCアナリティクス室室長)