

チームデータ分析における機械学習の活用手法と その適用事例

○仲間 大輔、入江 崇介、宇野 渉

株式会社リクルートマネジメントソリューションズ HR Analytics & Technology Lab

1 問題

企業の保有する人事・組織データを用いたチーム分析をマネジメント支援につなげていくことの実務的なニーズは高いものの、そうした分析は必ずしも一般的になっていない。その理由の1つとして、チームに関する人事・組織データの「横長」構造の問題が挙げられる。

チーム分析においてはチーム状態を個票データの集約として取り扱うことが多いが、その集約の方法が複数に及ぶため[e.g., 1]、チームレベルの変数が膨大になる。それに対して、一般にチーム数はそれほど多くはならないことが多く、変数数の相対的な多さが探索的な分析を進める際の制約となっていると言える。

そこで、本研究では、機械学習的手法を活用して、チーム分析におけるこうした制約を解消することを試みた。チーム分析では単一の変数の直接的な効果はそれほど頑健にはならず調整変数の考慮が重要となるというチーム研究の知見を踏まえ[2]、以下のモデルを想定する。

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_1 X_2$$

具体的な手法としては、まず、成果変数と関連度の高い説明変数を線形相関と Boruta（機械学習モデルと検定を用いた変数選択手法）によって抽出し、次に、そこで見られた影響関係を調整する変数をブートストラップ法を用いながら探索する、という2段階の手続きとした。つまり、上のモデルに適合する2つの変数を大量のチーム変数群から選択するために、機械学習的手法を段階的に活用した。モデルに用いる変数数を2つに限定するのは、分析結果をもとにしたチームマネジメント施策の実施を容易にすることを重視しているためである。

2 適用事例

本発表では、ある企業の一事業部門を対象に行った分析結果を報告する（チーム数=17）。分析には、各従業員の所属組織情報と3種類の組織サーベイ・アセスメントへの回答を統合したデータセットを用いた。個々人の回答を、平均、SD、最大値、最小値等によりチームごとに集約し、約1400のチームレベルの変数を分析対象とした。成果変数としては、チームの生産性についてのサーベイ項目のチーム平均を用いた。分析の詳細ならびに結果は大会当日に報告する。

[1] Barrick, Stewart, Neubert, & Mountet. (1998). Relating member ability and personality to work-team processes and team effectiveness. *Journal of Applied Psychology*, 83(3), 377-391.

[2] Mathieu, Hollenbeck, van Knippenberg, & Ilgen. (2017). A century of work teams in the *Journal of Applied Psychology*. *Journal of Applied Psychology*, 102(3), 452-467.