

## 概要書

1株当たり利益(Earnings Per Share: EPS)は、企業の当期純利益を発行済株式数で割ったものであり、規模に依存しない企業の収益性を捉えた指標である。企業外部のステークホルダーである投資家は、企業の将来のEPSの予測をもとに投資判断を行い、企業内部の経営者は、将来のEPSの予測を用いて営業予算の作成や設備投資の判断などの重要な意思決定を行う。したがって、EPSを正しく予測することは企業内外の幅広いステークホルダーにとって重要である。

企業のEPS予測は、人的な予測と、統計的・機械的な予測の2つに大別できる。人的な予測の代表として、証券アナリストが公表するアナリスト予測がある。他方、統計的・機械的な予測とは、過去の実績データをもとに何らかの時系列モデルを用いて将来のEPSを予測するものである。モデルに基づく予測は、予測値を導出するまでの過程を全て自動化できるため、人的な予測に比べてコストが低いという特徴を有している。従来、EPSを予測する時系列モデルとして、年次EPSについてはランダムウォーク、四半期EPSについてはBrown and Rozeff (1979)、Griffin (1977)、Foster (1977)の3つの自己回帰和分移動平均モデルがうまく描写するとされてきた。しかし、これらの時系列モデルによるEPS予測と、アナリストによる予測の精度を比較すると、アナリストの方が正確な予測を与えているのが現状である(桜井, 1990)。特に日本においては、時系列モデルによる予測よりもアナリスト予測の方が市場の期待利益として適切であると暗黙裡に見なされており、時系列モデルによるEPS予測の研究は、現在では衰退している(太田, 2006)。

上記の比較対象とされた時系列モデルは、あくまでも伝統的な単変量で線形の時系列モデルである。一方、近年のモデルによる四半期EPS予測の分野では、売掛金、棚卸資産、資本的支出といった将来の四半期EPSを予測する情報があるとされる会計変数(ファンダメンタル会計変数)を用いた多変量モデルがより精度の高い予測を与えるとされている。また、時系列モデルとして機械学習アルゴリズムを用いた研究も盛んである。機械学習アルゴリズムには、従来の統計的時系列モデルに比べて、高次元データの処理に長けたモデルや非線形性を捉えられるモデルが豊富にある(Cao and You, 2020)。特に、四半期EPSデータは(i)財務的、(ii)季節的、(iii)非線形な特徴を有しており(Hill et al., 1994)、機械学習アルゴリズムを用いることで、多変量なファンダメンタル会計変数から予測情報を抽出し、企業利益データの非線形性を捉え、精度の高い四半期EPS予測が得られると考えられる。

そこで、日本企業データにおいて、機械学習アルゴリズムは、従来の統計的時系列モデルに比べ四半期EPS予測の精度を向上させるのか、人的な予測よりも

高い精度の予測を与えるのかを検証することを本稿の目的とする。具体的には日本企業を対象に単変量線形モデル,多変量線形モデル,単変量機械学習アルゴリズム,多変量機械学習アルゴリズムによる四半期EPS予測を行い,時系列モデル間の予測精度比較,機械学習アルゴリズムによる予測とアナリスト予測の精度比較をそれぞれ実施する。なお,本稿では機械学習アルゴリズムとしてRidge回帰,LASSO回帰,Elastic Net回帰,ランダムフォレスト回帰,ニューラルネットワークを用いる。追加的な検証として,本稿で得られた四半期EPS予測値に基づいた株価予想利益率からロング・ショートポートフォリオを構築し,機械学習アルゴリズムによる四半期EPS予測の投資指標としての有用性を考察する。

本稿のサンプル期間は,日本の四半期報告制度が適用された2008年度(2008年4月1日)から現在2020年度(2021年3月31日)までの計52四半期とする。サンプル企業は,東京証券取引所一部上場企業の1,003社とする。また,予測期間は2018年度第1四半期~2020年度第4四半期の12四半期とし,予測のプロセスとしてはローリングサンプルを用いる。各企業に対して,長さ41期のローリングサンプルを12個作成し,同じ構造のモデルをローリングサンプルごとに推定することで,1企業あたりの12個の予測値系列を得る。そして得られた予測値系列をもとに,予測精度指標(平均絶対誤差,平均絶対誤差率,平均二乗誤差率)を各予測手法ごとに算出し,予測精度を比較する。最後に,2つの手法間の予測精度の差に統計的有意性があるかを検証するため,Diebold-Mariano検定を各企業に対して行う。

まず,機械学習アルゴリズムと伝統的な線形時系列モデルの予測精度を比較すると,単変量予測では予測精度に明らかな差は見られなかったが,多変量予測では,機械学習アルゴリズムの方が精度の高い予測を与えた。この結果から,機械学習アルゴリズムは,伝統的な線形時系列モデルでは捉えられない将来の四半期EPSとファンダメンタル会計変数の関係性を捉えることができ,機械学習アルゴリズムを用いることで日本企業のEPS予測のパフォーマンスを向上させる可能性が示唆される。次に,機械学習アルゴリズムとアナリストの予測精度を比較すると,前者は後者と同等もしくはそれ以上に高い精度の予測を与えることが明らかになった。企業利益予測の研究分野では,長らく時系列モデルによる予測よりもアナリスト予測の方が適切であると暗黙裡に見なされてきたが,本稿の結果はその認識を覆すものとなった。さらに本稿の追加的な検証から,機械学習アルゴリズムによる四半期EPS予測値は投資指標としても優れていることがわかった。

機械学習アルゴリズムはモデルの設定によって予測のパフォーマンスが大きく変動するため,本稿で得られた結果は必ずしも最良なものではないかもしれない。しかしながら,本稿の結果より,機械学習アルゴリズムはアナリストと比べて劣らない,もしくは優れた四半期EPS予測を与えることが明らかになった。このことから,現在日本で衰退してしまっている時系列モデルによる利益予測の研究が,機械学習アルゴリズムを用いることで再興する余地が十分にあることが示唆される。

## 参考文献

- Brown, Lawrence D and Michael S Rozeff (1979) “Univariate time-series models of quarterly accounting earnings per share: A proposed model,” *Journal of Accounting Research*, Vol. 17, No. 1, pp. 179–189.
- Cao, Kai and Haifeng You (2020) “Fundamental Analysis Via Machine Learning,” Technical report, Working Paper.
- Foster, George (1977) “Quarterly accounting data: Time-series properties and predictive-ability results,” *Accounting Review*, Vol. 52, No. 1, pp. 1–21.
- Griffin, Paul A (1977) “The time-series behavior of quarterly earnings: preliminary evidence,” *Journal of Accounting Research*, Vol. 15, No. 1, pp. 71–83.
- Hill, Tim, Leorey Marquez, Marcus O’Connor, and William Remus (1994) “Artificial neural network models for forecasting and decision making,” *International journal of forecasting*, Vol. 10, No. 1, pp. 5–15.
- 太田浩司 (2006) 「経営者予想に関する日米の研究: 文献サーベイ」, 『武蔵大学論集』, 第54巻, 第1号, 53–94頁.
- 桜井久勝 (1990) 「会計利益の時系列特性と利益予測」, 『経営学・会計学・商学研究年報 / 神戸大学大学院経営学研究科編』, 第36号, 45–98頁.