

論 文

機械学習と SHAP を用いた財務指標と株価変化の分析

—— 日本の自動車関連企業を対象として ——

山 口 健 二

概 要

本研究は、日本の自動車部門の会社 62 社を対象に、財務指標が株価に与える影響を機械学習で可視化することを行った。XGBoost による回帰モデルで、1 か月、1 年、5 年、10 年の株価リターンを目的変数とし、説明変数には 5 年平均と 10 年平均の主要 6 財務指標（売上高成長率、税前利益率、棚卸資産、固定資産、ROA、ROE）を用いた 8 ケースを設定した。モデルのブラックボックス性を補うため、協力ゲーム理論に基づく SHAP 値を算出し、それぞれの財務指標が株価に与える変化を定量的に比較した。その結果、全体的な考察としては、SG（売上高成長率）がいずれのパターンにおいても株価リターンに対して寄与が高く、その寄与の向きは若干ブレがあるものの、SG の値が高いと株価リターンも高くなる傾向が見られた。本手法は、多変量・非線形の企業データを扱う上で、伝統的統計モデルを補完しつつ、新しい手法として機械学習による分析を解釈可能な形で提示し、ひいては、企業個別の財務データ分析へと解釈を深めていくことができる。

1. はじめに

株価と企業の財務データは、その企業の健全性、成長性、収益性を示す重要な指標である。例えば、売上高は、企業が製品やサービスの販売から得た総収入であり、企業の基本的な収益性と市場でどの程度のシェアを持っているか見ることができる。従来から「製造業では在庫回転率が高い企業の方が、営業利益率が高い」と言われている。これは、効率的な在庫管理により、キャッシュフローの改善や保管コストの低減、柔軟な市場対応ができ、それにより営業利益率の向上に繋がるという考え方である。また、企業の業績が好調で、売上高が上がった場合、株価も上がると予想されるが、これは、投資家が、企業が成長していると認識し、将来的な収益増加の可能性があると見なされるため、投資家が株を購入することで株価を押し上げることによるものである。ただし、これには、営業利益や経常利益といった利益率にも着目しなければならず、また、企業が属する業界の状況も加味する必要がある。即ち、いくら売上高が好調だといっても、業界全体の売り上げが下降している場合は、株価を押し上げるには難しいと考えられる。これは、投資家が将来の成長見通しを重視するためである。また、それ以外にも、政治的なリスクなども考慮する必要がある。このような、定性的な要因が存在するものの、定量的な評価を避けるのは好ましくなく、株価と企業の財務データの関係をなんらかの手法で分析する研究は重要である。

特に、近年では、従来の統計学的な手法に加え、機械学習を活用した分析手法が積極的に導入されつつある。回帰分析などの従来の統計的手法では説明変数の多重共線性や非線形の関係性を適切に捉えることが難しい場合があるが、機械学習を用いた手法は、大量かつ複雑なデータを分析しパターンを認識することなどで、非線形な関係性をも捉える能力を持つため、様々な分野で注目されている。その中でも、ランダムフォレストや勾配ブースティング (Gradient Boosting) をベースとした XGBoost などの手法は、株価や財務指標など複雑な関係性をもつデータの分析に応用されている。これにより、予測精度の向上や、従来の統計手法では明示的に評価しにくかった変数間の影響度合いの可視化など、新たな知見を提供している。このように、近年の機械学習の発展は、従来の定量的分析を補完し、より精緻な企業分析を可能にしている。

本研究では、機械学習を用いて、企業の株価は、財務データのどの要因によって変動するのか分析を行った。また、機械学習による説明は、ブラックボックス化しがちであるが、SHAP 値を算出することで、各予測結果の理由を定量的に明らかにした。

2. 既存研究

ここでは、株価や財務データの予測に関する既存研究について、いくつか説明をする。

株価を決定する要因については、数多くの研究が行われている。ファイナンス関連の研究による決定要因としては、企業財務に関する指標、配当、企業規模、未分配利益、ROA、ROE、負債比率などが挙げられる[1, 2]。伊藤レポートでは、日本企業の ROE は世界平均の ROE よりも低く、日本企業の目標は ROE 8%の達成であるとしている[3]。このように、投資家は ROE を重要なファクターとして捉えている。また、マクロ経済の観点からは、長期金利と短期金利の乖離、インフレ率、鉱工業生産などの指標も関連要因として研究されている[4]。

株価リターンを予測するために、様々な時系列分析モデルが用いられてきた。中でも、株価の変動性は重要な研究テーマであり、ARCH モデル[5]をはじめとする多くの手法が開発されている。しかし、時系列モデルを予測に用いる場合、株価の変動性に影響を与える要因を特定する必要があるが、一般的にはそのような要因を特定することは困難である。例えば、1970 年代には石油危機、金融政策の変更、ブレトンウッズ体制の崩壊などが株式市場やマクロ経済の変動性を高めたが、どの要因が最も重要であったかの判断は研究者によって見解が分かれる。

株価変動の研究には、株価下落のパターンと投資家の行動との関係を説明することに着目するものがある。例えば、天災・人災を問わず何らかの危機が発生すると、投資家（自動売買プログラムも含む）が反応し、経済指標や企業業績の指標を参考にしながら株式の売買を決定する。したがって、なんらかの危機が株価変動に及ぼす影響に関する研究は、投資家の動機や行動に及ぼす影響を測定して、市場全体への影響を分析するものと考えられる。

また、企業の財務パフォーマンスに影響を与える要因として、製品開発能力 (Product Development Capability : PDC) とサプライチェーン能力 (Supply Chain Capability : SCC) の相互作用を分析した研究もある。森田らは、製品開発とサプライチェーン管理を統合的に調整する能力が、企業が変化する競争環境に動的に適応し、持続的に高いパフォーマンスを実現するための重要な要素であることを明らかにした[6]。従来の研究では、製品特性とサプライチェーン特性の「適合 (fit)」が個別の状況に固定的に対応することが議論されてきたが、競争環境が時間とともに変化するため、この適合を動的に調整

できる能力の必要性が十分に検討されていなかった。そこで森田らは、「絶対的サプライチェーン戦略志向（Absolute Supply Chain Orientation Strategy：ASCOS）」という概念を導入し、製品開発能力とサプライチェーン能力の双方を強化・調整することにより、企業が長期的な競争優位性を維持できると仮説を立てた。そして、14 か国、303 工場から収集した国際的なデータベースを用いて統計分析を行った結果、製品開発能力（PDC）とサプライチェーン能力（SCC）は相互補完的な関係にあり、両能力を高いレベルで維持する企業は、製品開発の迅速性、オンタイムでの製品導入、生産量や製品ミックス変更に対する柔軟性において優れていることを明らかにした。さらに、多くの企業が高いパフォーマンスを達成するためには、製品開発能力よりもサプライチェーン能力の強化を相対的に優先すべきであるという示唆も示されている。

ここで述べたように、株価要因研究は多岐にわたるものの、機械学習を用いて非線形の寄与を精緻に定量化し、かつ産業内で企業ごとに異なる財務指標と株価の関係を SHAP で可視化する試みは限られている。

特に、売上成長・在庫回転・設備投資・利益率といった重要な財務指標が株価に及ぼす影響が、それらの分析期間の長短によって、重要度が入れ替わるかは十分に検証されていない。

そこで次章では、日本の自動車関連 62 社を対象に株価リターンと 5 年・10 年平均の財務データを組み合わせた 8 ケースの分析フレームワークを提示し、実証分析をするためのデータ構成と変数を設定する。

3. 分析データ

本研究では、日本企業の製造業の自動車部門の会社に対して、株価と財務データとの間にどのような関係があるかを調べた。本研究で用いたデータは次のとおりである。日経 NEEDS Financial QUEST（以下 NEED）で企業の株価および財務データを取得した。対象企業は、NEEDS の株式区分が親会社、業種分類が日経業種分類で製造業の自動車に属する企業である。今回の目的変数を算出する期間にデータがない企業や決算が 3 月末日でない企業を除外したところ、自動車に属する企業は、自動車会社 9 社、自動車部品会社 49 社、自動車車体会社 4 社の計 62 社となった。これらの企業に対して分析を行った。

目的変数を 4 通り、説明変数を 2 通り用意し、それぞれを組み合わせ、計 8 通りで分析をした。

目的変数は、以下のとおりである。例えば、1. の場合、2022 年 2 月の株価が 1000 で、2022 年 3 月の株価が 1200 の場合の株価リターンは 1.2 である。

1. 短期的株価変動を見るための 2022 年 2 月の終値を 1 としたときの 2022 年 3 月の株価リターン
2. 中期的株価変動を見るための、2021 年 3 月の終値を 1 としたときの、2022 年 3 月の株価リターン
3. 長期的株価変動を見るための、2017 年 3 月の終値を 1 としたときの、2022 年 3 月の株価リターン
4. 超長期的株価変動を見るための、2012 年 3 月の終値を 1 としたときの、2022 年 3 月の株価リターン

説明変数は、以下の通りである。

1. 2018 年から 2022 年の 5 年間の財務データ 6 項目の各平均値
2. 2013 年から 2022 年の 10 年間の財務データ 6 項目の各平均値

財務データの 6 項目は、以下のとおりである (表 1)。ただし、「(前年比) 売上高・営業収益 [累計]」は平均値として幾何平均を使用し、「売上高成長率 (Sales Growth)」の指標とした。分析上、表 1 にあるように用語を英語で置き換えている。また、目的変数と説明変数を組み合わせて、それぞれ T1-Ave5, T1-Ave10, T2-Ave5, T2-Ave10, T3-Ave5, T3-Ave10, T4-Ave5, T4-Ave10 と名称をつけて計 8 ケースとした (表 2)。

分析手法は、機械学習回帰とその予測結果を説明する SHAP 値による分析を行った。以前、上記のケースのうち T1-Ave5, T1-Ave10, T2-Ave5, T2-Ave10 に対して、重回帰分析を行った [7] が、重相関 R、重決定 R² ともに、数値が大きくなかったため、今回は機械学習回帰と SHAP 値による分析を中心に行った。

表 1：説明変数として使用した財務データと分析上の名称（5 年間の財務データの平均値の場合）

日経 NEEDS Financial QUEST 上の名称	分析上の名称 (重回帰分析)	分析上の名称 (SHAP 値による分析)
(前年比) 売上高・営業収益 [累計]	(Ave5) Net Sales	SHAP-(Ave5)SG
(売上高に対する) 税金等調整前当期純利益 [累計]	(Ave5) Net Income before Income Taxes and Others / Net Sales	SHAP-(Ave5)NI
(売上高に対する) 棚卸資産	(Ave5) Inventories / Net Sales	SHAP-(Ave5)IVT
(売上高に対する) 有形固定資産	(Ave5) Tangible Fixed Assets / Net Sales	SHAP-(Ave5)TFA
使用総資本事業利益率	(Ave5) ROA	SHAP-(Ave5)ROA
ROE (自己資本利益率) (A)	(Ave5) ROE	SHAP-(Ave5)ROE

(10 年間の財務データの平均値を使用した場合は、(Ave10) となる)

表 2：目的変数と説明の変数の組み合わせと各ケース

ケース	T1-Ave5	T1-Ave10	T2-Ave5	T2-Ave10	T3-Ave5	T3-Ave10	T4-Ave5	T4-Ave10
目的変数	(短期的株価変動) 2022 年 2 月の終値を 1 としたときの 2022 年 3 月の株価リターン		(中期的株価変動) 2021 年 3 月の終値を 1 としたときの 2022 年 3 月の株価リターン		(長期的株価変動) 2017 年 3 月の終値を 1 としたときの 2022 年 3 月の株価リターン		(超長期的株価変動) 2012 年 3 月の終値を 1 としたときの 2022 年 3 月の株価リターン	
説明変数	2018 年から 2022 年の 5 年間の財務データ 6 項目の各平均値	2013 年から 2022 年の 10 年間の財務データ 6 項目の各平均値	ケース T1-Ave5 と同じ	ケース T1-Ave10 と同じ	ケース T1-Ave5 と同じ	ケース T1-Ave10 と同じ	ケース T1-Ave5 と同じ	ケース T1-Ave10 と同じ
対象企業	自動車メーカー 9 社 自動車部品会社 49 社 自動車車体会社 4 社 計 62 社							

4. 機械学習回帰と SHAP 値による分析手法について

本研究では、まず機械学習による回帰を行い、その結果から SHAP 値 (SHapley Additive exPlanations) を求め、どの財務データが株価に影響を与えているのかを考察する。機械学習回帰の手法として、XGBoost を使用した。XGBoost (eXtreme Gradient Boosting) は機械学習回帰で用いられるアルゴリズムの一つで、決定木をベースにしており、勾配ブースティングという手法を用いて、複数の決定木を組み合わせることで回帰予測を行う。勾配ブースティングは、前の決定木の誤差を次の決定木が修正することで、モデルの性能を向上させる手法である。

また、SHAP 値は、各企業の目的変数（株価リターン）に対する予測を、各説明変数の寄与に分解したものである。SHAP 値が正の場合、その説明変数はモデルの予測値をベースライン（全体平均予測）より押し上げる方向に寄与したことを意味し、逆に負の場合は予測値を押し下げる方向に寄与したことを意味する。したがって、SHAP 値を用いることで、例えば、「ある企業では ROE の水準が株価リターン予測を上昇方向に強く寄与している」といった解釈が可能となる。さらに、ある説明変数について、各企業における各説明変数の SHAP 値と株価リターンをプロットすることで、寄与のばらつきや企業ごとの特性を視覚的に把握できる。なお、機械学習の分野では、XGBoost の入力変数を「特徴量」と呼ぶことが多いが、本研究では、「説明変数」と呼ぶことにする。

全体的な傾向を見るのであれば、SHAP 値の要約プロットを参照することで、目的変数に対する各説明変数の寄与度を把握できる（図 1、図 3、図 5、図 7、図 9、図 11、図 13、図 15）。具体的には、SHAP 値の絶対値の平均値を取ることで、今回分析した自動車産業に属する企業において、各説明変数が株価リターンにどの程度寄与しているかを測定できる。ここで注意しなければならないのは、SHAP 値の要約プロットの寄与度は、SHAP 値の絶対値の平均を取っているため、説明変数が株価リターンに対してプラス方向／マイナス方向にどのように寄与したかは、ビースウォームプロットで確認する必要がある。

ビースウォームプロットは、各企業における各説明変数の SHAP 値（寄与度）を可視化したものである（図 2、図 4、図 6、図 8、図 10、図 12、図 14、図 16）。横軸は説明変数が予測に与える寄与度（SHAP 値）、縦軸は説明変数の種類を表す。点の色は各説明変数の元の値を示し、一般に小さい値が青、大きい値が赤として描画される。この図により、説明変数の値と予測への寄与の関係を直感的に把握できる。例えば、負の寄与（左側）に青い点が集中し、正の寄与（右側）に赤い点が集中している場合、その説明変数の値が大きい（赤い）ほど目的変数に正の影響を及ぼすことを意味する。逆に、負の寄与（左側）に赤い点が集中し、正の寄与（右側）側に青い点が多く分布している場合は、その説明変数の値が小さい（青い）ほど目的変数に正の影響を与える傾向を示す。さらに、点の分布が横方向に広い説明変数の種類ほど、サンプル間で寄与度が大きく変化しており、モデル予測に対して相対的に重要であると解釈できる。一方、点が中央付近に集中している説明変数の種類は、予測への影響が限定的であることを示唆する。

5. 分析結果と考察

ここでは、各ケースにおいて、どの財務指標が影響を与えているか考察するとともに、SHAP 値の要約プロットとビースウォームプロットを掲載する。

本研究の分析は、対象を自動車に属する企業 62 社に限定したデータに基づいており、結果の一般化には限界がある。しかしながら、従来の相関分析や線形回帰とは異なり、XGBoost と SHAP を組み合わせることで、財務指標と短期的株価リターンとの非線形かつ複雑な関係を部分的に可視化できる点に方法的意義がある。特に SHAP 値を用いることにより、各財務指標が株価変動に寄与する方向性と大きさを同時に把握できるため、個別の説明変数の影響を直感的に解釈することが可能となる。本研究では、T1-Ave5 から T4-Ave10 までの 8 ケースについて、SHAP 要約プロットおよびビースウォームプロットを用いた考察を順次提示する。

T1-Ave5:

2022 年 2 月の終値を基準とした 1 か月後の株価リターンを目的変数とし、説明変数として過去 5 年間の財務指標平均を用いたものである。

SHAP 値の要約プロット (図 1) を見ると、(Ave5)ROA が最も大きな寄与を示し、次いで (Ave5)SG (売上高成長率) が続いた。また、ビースウォームプロット (図 2) を見ると、(Ave5)ROA の値が小さい企業において、株価リターンに対して正の寄与を示す傾向が確認された。

図 1 : T1-Ave5 での SHAP 値の要約プロット

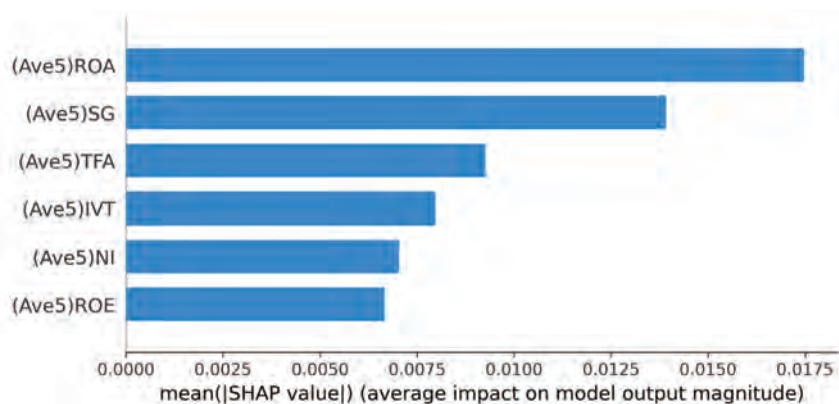
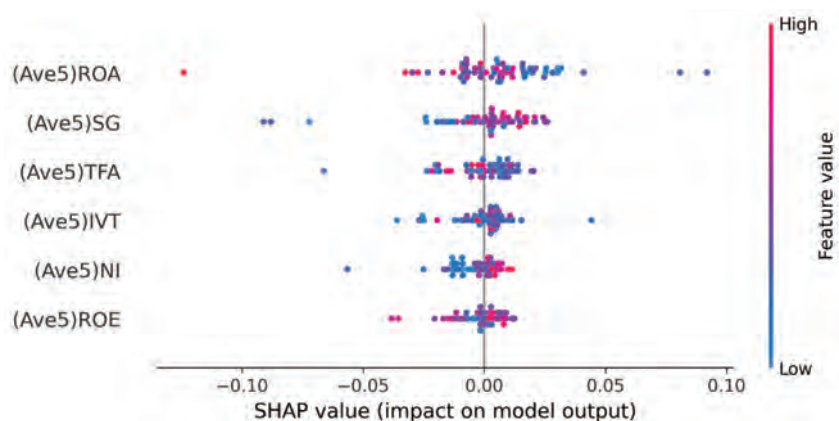


図 2 : T1-Ave5 でのビースウォームプロット



T1-Ave10：

2022 年 2 月の終値を基準とした 1 か月後の株価リターンを目的変数とし、説明変数として過去 10 年間の財務指標平均を用いたものである。

SHAP 値の要約プロット（図 3）を見ると、(Ave10)SG（売上高成長率）が最も大きな寄与を示し、次いで (Ave10)ROA、(Ave10)ROE が続いた。これは、直近 10 年間の平均的な売上高成長率や収益性指標が、2022 年 2 月を基準とした 1 か月後の株価リターンに対して相対的に大きな説明力を持つことを示している。また、ビースウォームプロット（図 4）を見ると、(Ave10)SG では、値が大きい場合（赤い点）が右側（正の SHAP 値）に分布する傾向が見られ、長期的に売上高成長率の高い企業は短期的な株価リターンにプラスの影響を与える可能性が示唆される。一方、(Ave10)ROA については、値が小さい場合（青い点）が右側に分布する傾向がみられ、T1-Ave5 と同様に「ROA が低い企業ほど短期的株価リターンに正の寄与を与える」傾向が確認された。

図 3：T1-Ave10 での SHAP 値の要約プロット

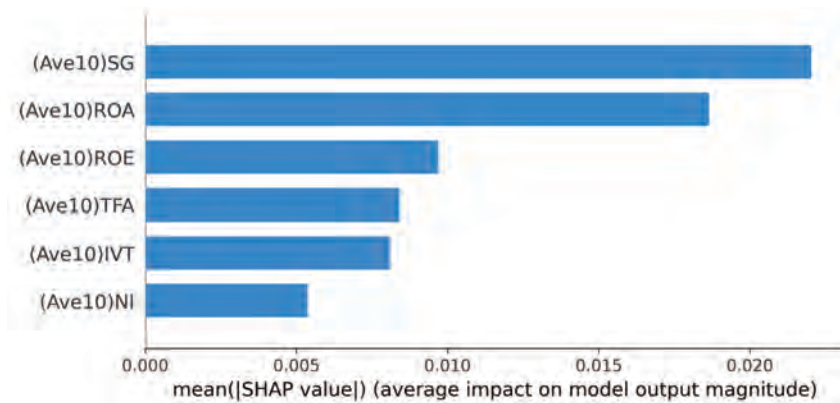
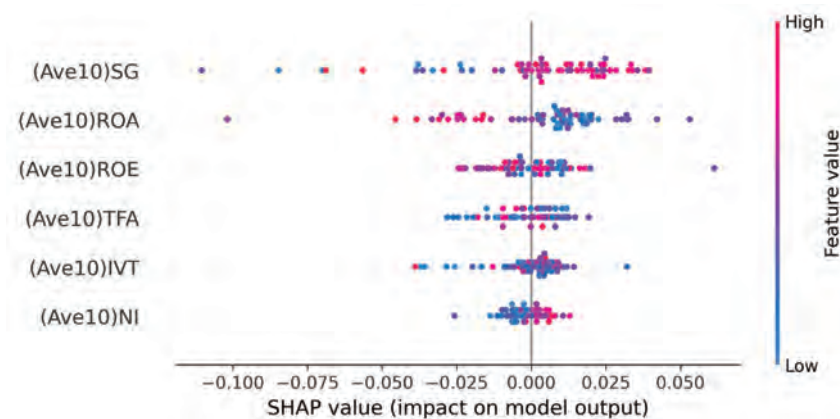


図 4：T1-Ave10 でのビースウォームプロット



T2-Ave5:

2021 年 3 月の終値を基準とした 1 年後の株価リターンを目的変数とし、説明変数として過去 5 年間の財務指標平均を用いたものである。

SHAP 値の要約プロット (図 5) を見ると、(Ave5)ROE (自己資本利益率) が最も大きな寄与を示し、次いで (Ave5)SG (売上高成長率)、(Ave5)NI (当期純利益) が続いた。これは、過去 5 年間の平均的な収益性や成長性指標が比較的大きな説明力を持つことを意味する。また、ビースウォームプロット (図 6) を見ると、(Ave5)ROE では値が高い場合 (赤い点) が左側 (負の SHAP 値) に分布する傾向がみられ、収益性の高い企業ほど短期的株価リターンにマイナスの寄与を与える可能性が示唆された。一方、(Ave5)SG では、値が大きき場合 (赤い点) が左側 3 個ほど確認できるが、それを除けば全体的に右側に分布しており、売上高成長率の高い企業が株価リターンに正の影響を与える傾向が確認された。

図 5 : T2-Ave5 での SHAP 値の要約プロット

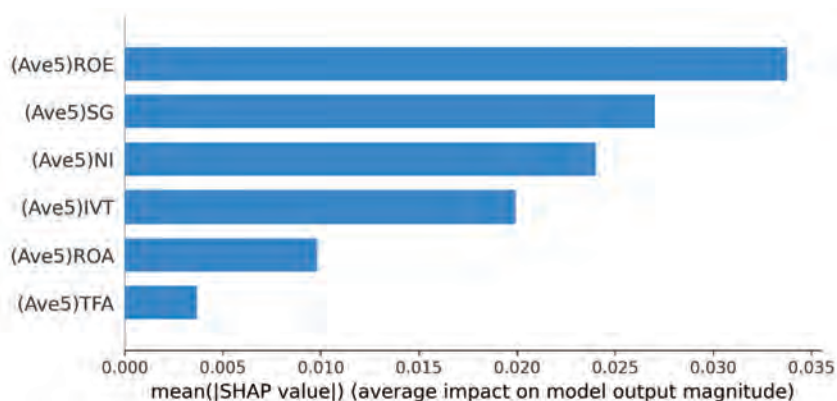
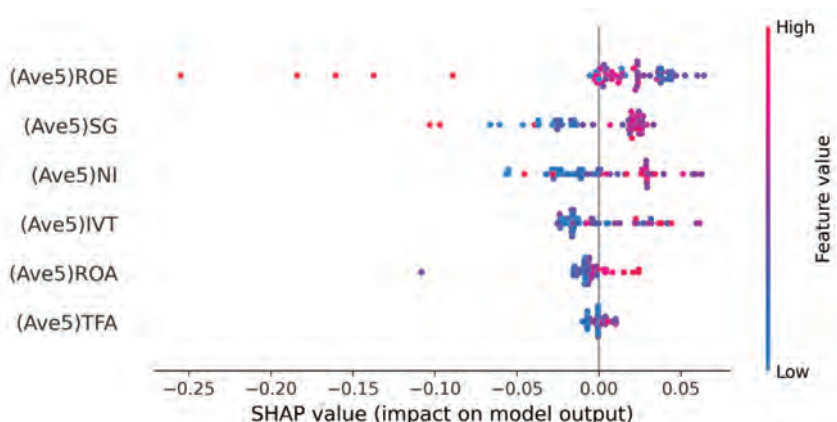


図 6 : T2-Ave5 でのビースウォームプロット



T2-Ave10:

2021 年 3 月の終値を基準とした 1 年後の株価リターンを目的変数とし、説明変数として過去 10 年間の財務指標平均を用いたものである。

SHAP 値の要約プロット（図 7）を見ると、(Ave10)NI（当期純利益）が最も大きな寄与を示し、次いで (Ave10)SG（売上高成長率）、(Ave10)IVT（棚卸資産）が続いた。また、ビースウォームプロット（図 8）を見ると、(Ave10)NI では説明変数の値が大きい（赤い点）が左右に分布する傾向が見られ、企業によって、当期純利益が大きいことが株価リターンを押し下げる方向・押し上げる方向のどちらかに寄与することが示唆される。

図 7：T2-Ave10 での SHAP 値の要約プロット

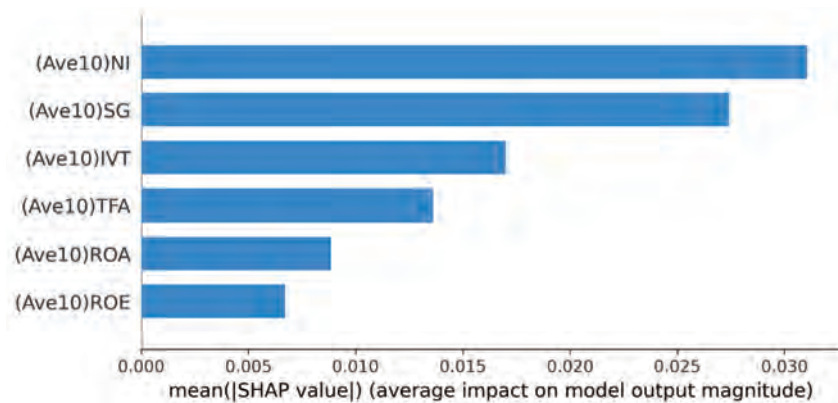
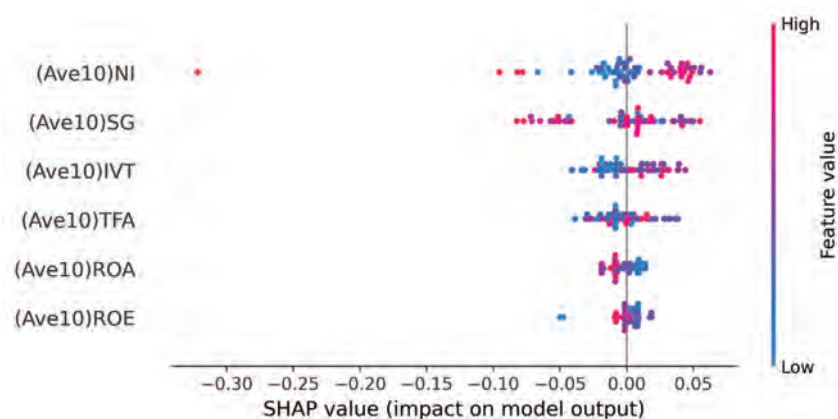


図 8：T2-Ave10 でのビースウォームプロット

**T3-Ave5:**

2017 年 3 月の終値を基準とした 5 年後の株価リターンを目的変数とし、説明変数として過去 5 年間の財務指標平均を用いたものである。

SHAP 値の要約プロット（図 9）を見ると、(Ave5)SG（売上高成長率）が最も大きな寄与を示し、

次いで (Ave5)NI (当期純利益), (Ave5)TFA (有形固定資産) が続いた。また, ビースウォームプロット (図 10) を見ると, (Ave5)SG については, 左側に紫の点が若干あるが, 値が大きい場合 (赤い点) が右側 (正の SHAP 値) に分布する傾向が見られる。

図 9 : T3-Ave5 での SHAP 値の要約プロット

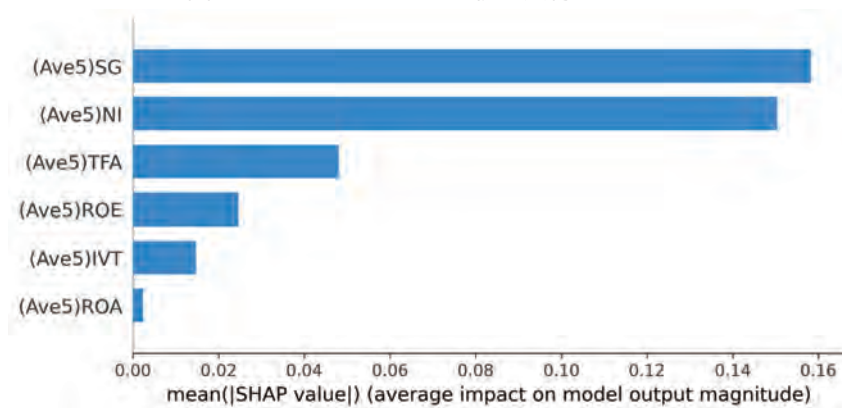
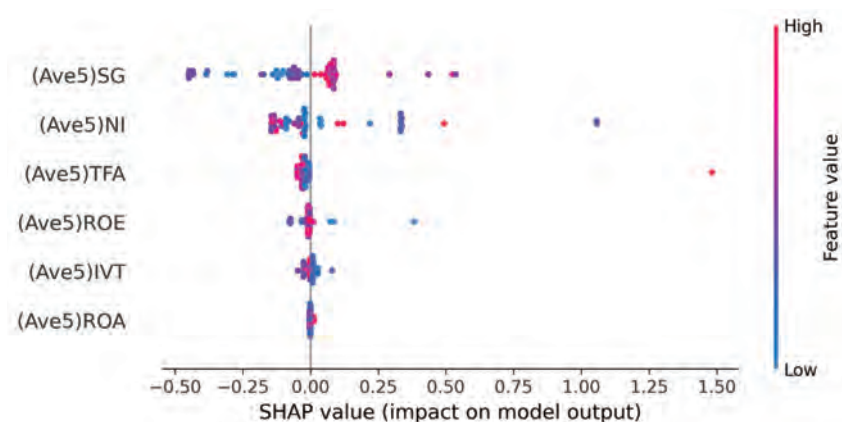


図 10 : T3-Ave5 でのビースウォームプロット



T3-Ave10 :

2017 年 3 月の終値を基準とした 5 年後の株価リターンを目的変数とし, 説明変数として過去 10 年間の財務指標平均を用いたものである。

SHAP 値の要約プロット (図 11) を見ると, (Ave10)IVT (棚卸資産) が最も大きな寄与を示し, 次いで (Ave10)NI (当期純利益), (Ave10)SG (売上高成長率) が続いた。また, ビースウォームプロット (図 12) を見ると, (Ave10)IVT については値が高い場合 (赤い点) が正の SHAP 値側に分布しており, 棚卸資産が高い企業ほど株価リターンにプラスの寄与を与える傾向が確認された。(Ave10)NI についても同様の傾向が見られた。

図 11：T3-Ave10 での SHAP 値の要約プロット

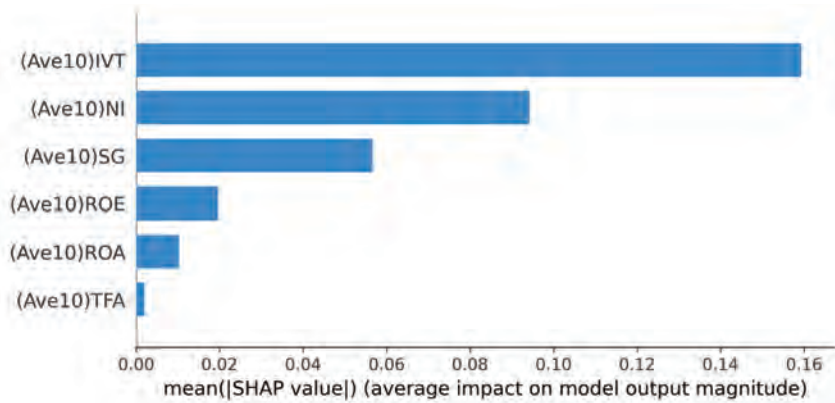
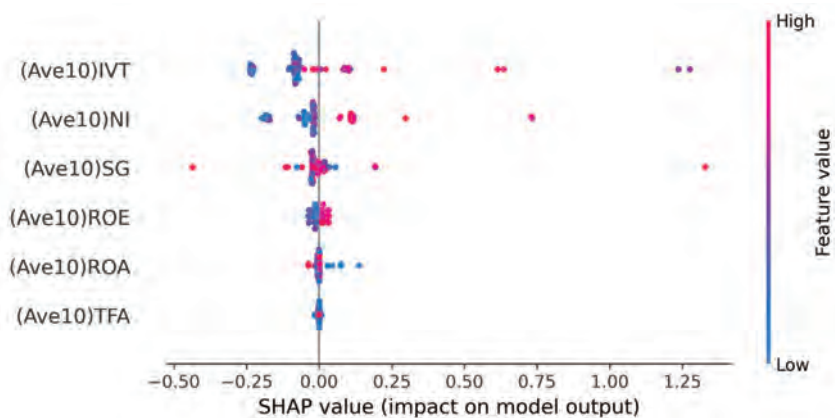


図 12：T3-Ave10 でのビースウォームプロット



T4-Ave5:

2012 年 3 月の終値を基準とした 10 年後の株価リターンを目的変数とし，説明変数として過去 5 年間の財務指標平均を用いたものである。

SHAP 値の要約プロット（図 13）を見ると，(Ave5)SG（売上高成長率）が最も大きな寄与を示し，次いで (Ave5)TFA（有形固定資産），(Ave5)IVT（棚卸資産）が続いた。また，ビースウォームプロット（図 14）を見ると，(Ave5)SG については値が高い場合（赤い点）が正の SHAP 値側に分布しており，売上高成長率が高い企業ほど株価リターンにプラスの寄与を与える傾向が確認された。また，(Ave5)TFA については，説明変数の値が高い場合（赤い点）が正負両方向に分布しており，売上高成長率が高い企業でも株価リターンへの寄与は一樣ではなく，プラスとマイナスの両面が存在することが分かる。

図 13 : T4-Ave5 での SHAP 値の要約プロット

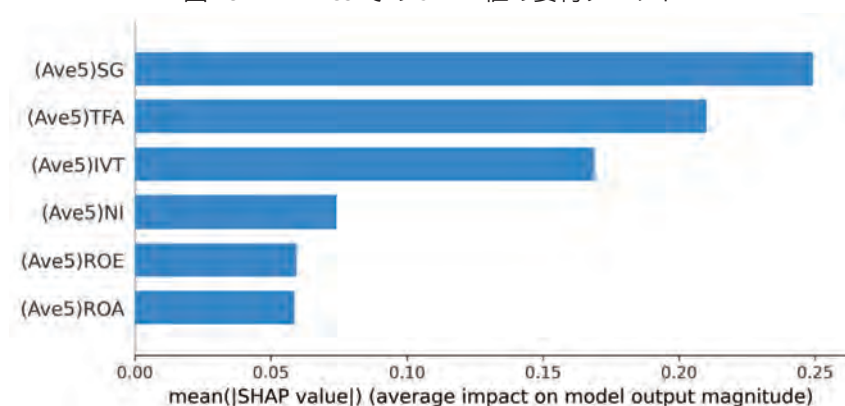
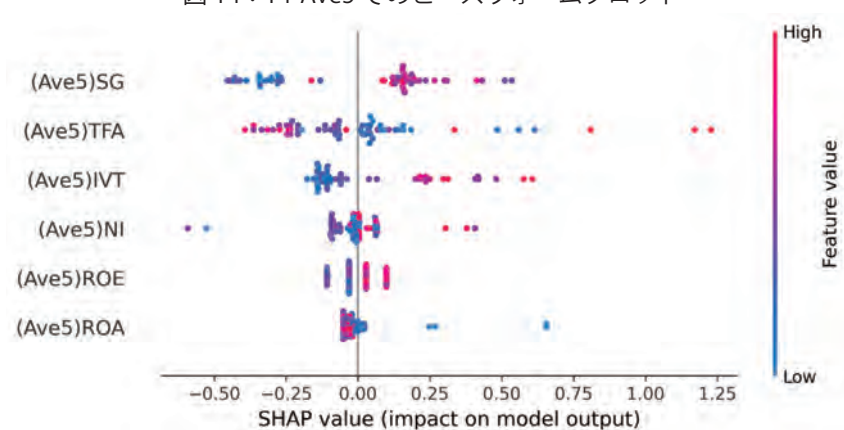


図 14 : T4-Ave5 でのビースウォームプロット



T4-Ave10 :

2012 年 3 月の終値を基準とした 10 年後の株価リターンを目的変数とし、説明変数として過去 10 年間の財務指標平均を用いたものである。SHAP 値の要約プロット（図 15）を見ると最も影響度が高いのは（Ave10）IVT（棚卸資産）であり、株価リターンに対して強い説明力を持つ。次いで（Ave10）NI（当期純利益）、（Ave10）SG（売上高成長率）が続いている。ビースウォームプロット（図 16）を見ると、（Ave10）IVT については値が高い場合（赤い点）が正の SHAP 値側に分布しており、棚卸資産が高い企業ほど株価リターンにプラスの寄与を与える傾向が確認された。（Ave10）NI、（Ave10）SG についても同様の傾向が見られた。

図 15：T4-Ave10 での SHAP 値の要約プロット

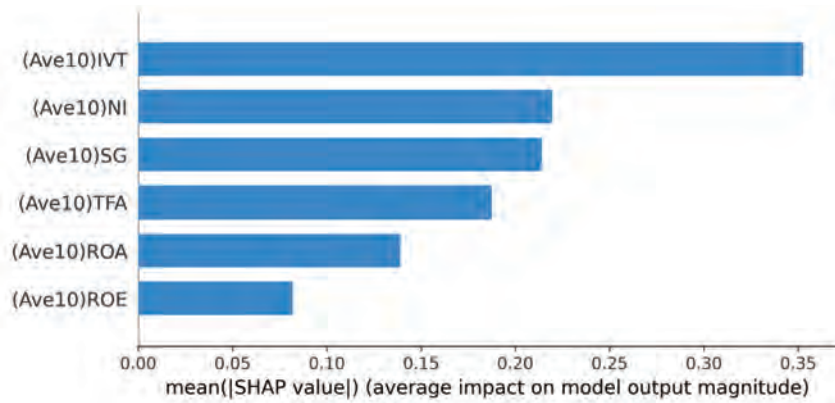
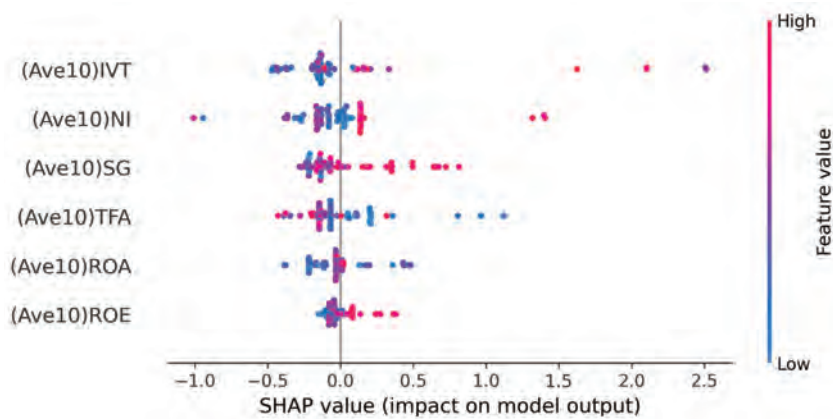


図 16：T4-Ave10 でのビースウォームプロット



これらの 8 ケースの結果を俯瞰するために、各パターンの SHAP 値の要約プロット（図 1，図 3，図 5，図 7，図 9，図 11，図 13，図 15）の寄与度を順位でまとめた（表 3）。

表 3：各パターンの SHAP 値の要約プロットの寄与度の順位

	T1-Ave5	T1-Ave10	T2-Ave5	T2-Ave10	T3-Ave5	T3-Ave10	T4-Ave5	T4-Ave10
1 位	ROA	SG	ROE	NI	SG	IVT	SG	IVT
2 位	SG	ROA	SG	SG	NI	NI	TFA	NI
3 位	TFA	ROE	NI	IVT	TFA	SG	IVT	SG
4 位	IVT	TFA	IVT	TFA	ROE	ROE	NI	TFA
5 位	NI	IVT	ROA	ROA	IVT	ROA	ROE	ROA
6 位	ROE	NI	TFA	ROE	ROA	TFA	ROA	ROE

株価リターンに寄与する説明変数は、「株価リターンを測るための期間」と「財務指標の集計期間」の掛け合わせで系統的に入れ替わる。

例えば、株価リターンを測るための期間が、1ヶ月間の短期的株価変動（T1）では、ROAとSGが突出しており、いまだに「資産を効率良く回し」、「成長しているか」が関係している。一方で、株価リターンを測るための期間が、1年間、5年間、10年間の中期的、長期的、超長期的株価変動（T2、T3、T4）では、SGとNI、IVTが上位にあり、売上高成長率と当期純利益、そして棚卸資産が、株価リターンに寄与していることがわかった。その中でも、IVTは財務指標の集計期間が10年間（T2-Ave10、T3-Ave10、T4-Ave10）の場合、5年間（T2-Ave5、T3-Ave5、T4-Ave5）よりも株価リターンに与える影響が強くなっている。すなわち、財務指標の集計期間が長いと棚卸資産が、より株価リターンに影響を与えることが示唆された。

「4. 機械学習回帰とSHAP値による分析手法について」で説明したように、SHAP値の要約プロットの寄与度は、SHAP値の絶対値の平均を取っているため、説明変数が株価リターンに対してプラス方向／マイナス方向にどのように寄与したかは、ビースウォームプロットで確認する必要があった。したがって、表3を俯瞰することで株価リターンへの影響が強いと確認された説明変数についてのみ、ビースウォームプロット（図2、図4、図6、図8、図10、図12、図14、図16）の結果をもとに、説明変数が株価リターンに対して、どのように寄与したかを詳しく俯瞰するため、表4にまとめた。

- ・「青 赤」は、左側に青い点、右側に赤い点が多く分布しており、説明変数の値が大きいと、株価リターンに対して正の寄与を示す傾向があるもの
- ・「赤 青」は、左側に赤い点、右側に青い点が多く分布しており、説明変数の値が小さいと、株価リターンに対して正の寄与を示す傾向があるもの
- ・「赤青赤」は、説明変数の値が大きい（赤い点）がビースウォームプロットの左右両方向に分布しており、説明変数の値が大きいと、株価リターンに対して正の寄与または負の寄与のどちらかを示す傾向があるもの
- ・「ランダム」は、説明変数の値の大きさと株価リターンに対する寄与がランダムとみられるもの

この表4から、以下のことが言える。まず、株価リターンを測るための期間が、1ヶ月間の短期的株価変動（T1）の場合、ROAは赤 青、SGは青 赤である。したがって、ROA値が小さい企業やSGの値が大きい企業において、株価リターンに対して正の寄与を示す傾向がある。

表4：表3の一部に、ビースウォームプロットの点の色の分布を記載

	T1-Ave5	T1-Ave10	T2-Ave5	T2-Ave10	T3-Ave5	T3-Ave10	T4-Ave5	T4-Ave10
1位	<u>赤 青</u>	<u>青 赤</u>		赤青赤	<u>青 赤</u>	<u>青 赤</u>	<u>青 赤</u>	<u>青 赤</u>
2位	<u>青 赤</u>	<u>赤 青</u>	赤青赤	ランダム	赤青赤	<u>青 赤</u>		<u>青 赤</u>
3位			<u>青 赤</u>	<u>青 赤</u>		ランダム	<u>青 赤</u>	<u>青 赤</u>
4位								
5位								
6位								

一方で、株価リターンを測るための期間が、1 年間、5 年間、10 年間の中期的、長期的、超長期的株価変動（T2, T3, T4）の場合、SG や NI の値が大きい企業において、株価リターンに対して正の寄与を示す傾向がある。ただし、「赤青赤」が SG は 1 つ、NI は 2 つあるので、企業によっては、SG や NI の値が大きいと株価リターンに対して負の寄与を示すこともある。

そして、財務指標の集計期間が 10 年間（T2-Ave10, T3-Ave10, T4-Ave10）の場合、SG と NI と IVT が影響を与えるトップ 3 であったが、このときの IVT を見ると、「青 赤」であり、IVT の値が大きい企業において、株価リターンに対して正の寄与を示す傾向があることが分かる。

また、いずれのパターンにおいても、SG は、上位 3 位以内に必ず入っており、重要な指標であることが明らかになった。

6. まとめと今後の課題

本研究は、日本の自動車関連 62 社を対象に、XGBoost と SHAP 値を用いて株価変動と財務指標の関係を 8 ケースにわたり可視化した。全体的な考察としては、SG（売上高成長率）がいずれのパターンにおいても株価リターンに対して寄与が高く、その寄与の向きは若干ブレがあるものの、SG の値が高いと株価リターンも高くなる傾向が見られた。

また、分析の結果、「株価リターンを測る期間」と「財務指標の集計期間」の二つの軸の組み合わせによって、寄与する説明変数の傾向が系統的に変化することを示した。すなわち、短期的な株価変動を説明する際には、ROA や SG（売上高成長率）といった収益性や成長性に直結する指標が強く作用し、企業の効率性や成長ポテンシャルが投資家の即時的な判断に影響していることが分かる。一方で、中期から長期、超長期にかけては、SG や NI（当期純利益）とともに IVT（棚卸資産）といった資産構造に関連する指標が上位に現れ、企業の持続的な売上成長や資産の活用度合いが株価リターンに大きく寄与していることが示唆された。

さらに、財務指標の集計期間が 5 年間か 10 年間かによっても結果は異なる。5 年間の集計では、直近の収益性や成長性が比較的強調されやすいのに対し、10 年間の集計では、企業の長期的な投資活動や資産蓄積が評価に反映されやすく、中でも IVT が上位を占める傾向が確認できた。つまり、財務データの集計期間が長くなるほど「短期的な業績変動」よりも「長期的な成長・資産戦略」が株価にとってより重要な要素となる。

本研究の分析結果は、単に「どの変数が重要か」を示すだけではなく、株価を動かす要因の構造が、どの時間スケールで分析するかによって大きく変わることを明らかにしている。つまり、短期投資を前提とする場合と、長期投資を前提とする場合とでは、投資判断において重視すべき財務指標が異なる可能性がある。今回の方法は、そうした「時間軸の違いによる変化」を整理して示すことができる点に大きな意義がある。

他方、課題も残る。まず、本研究は日本の自動車関連企業に限定した分析であり、国や業種によっては産業特性が異なるため注意が必要である。また、モデルの精度評価は学習データ内で完結しており、この期間を変えて検証することで、今回の結果をより補強することができる。さらに、今回は個別の企業の SHAP 値の考察までは至らなかったが、これは今後課題としたい。

これらを踏まえ、今後は、他業種・海外企業への拡張、予測精度と説明力を同時に高めるための検証とモデル設計、企業個別の経営意思決定を支援するダッシュボード開発などを進めることで、企業経済

データ分析へ機械学習と SHAP 値による解釈可能性を応用することが可能となる.

参考文献

- [1] 松井久勝, 音川和久, 会計情報のファンダメンタル分析, 中央経済社, 2013
- [2] 永島正康, グローバル・サプライチェーンにおける新しい製販協働のかたち, 2021
- [3] 経済産業省, 持続的成長への競争力とインセンティブ〜企業と投資家の望ましい関係構築〜, プロジェクト最終報告書 (伊藤レポート), 2014
- [4] 松井美樹, サプライチェーン・マネジメント, 放送大学教育振興会, 2021
- [5] 渡部敏明, 佐々木浩, ARCH 型モデルと Realized Volatility によるボラティリティ予測と Value-at-Risk, 日本銀行金融研究所, Discussion Paper No.2006-J-13, 2006
- [6] Michiya Morita, Jose A.D. Machuca, Jose L. Perez Díez de los Ríos, "Integration of product development capability and supply chain capability: The driver for high performance adaptation", International Journal of Production Economics 200 (2018) 68-82
- [7] 山口健二, 機械学習を用いた企業の特性分析, 日本大学経済学部経済科学研究所, ワーキングペーパー, No.23-01, 2024 年 3 月