

早稲田商学学生懸賞論文 要旨

※以下に、表題、要旨（1行35字・1000字以内）を記述してください。

日本において、四半期 GDP は該当四半期終了から約 45 日後にその 1 次速報値が発表される。こうした、四半期終了とその四半期の速報値発表のタイミングの差が、経済政策のラグを生じさせうる。一方で、政策担当者は目下の経済状況を反映した政策を速やかに立案する必要がある。経済の現況を予測する「ナウキャスト」という考え方は、現況に合った経済政策を即時的に立案可能にするという点で重要であることが知られている。本稿は、日本の 2012 年第 1 四半期から 2019 年第 2 四半期までの GDP 四半期成長率について、機械学習手法を用いたナウキャストを行った。説明変数には、主として GDP と相関が高いとされるマクロ経済指標を用いた。マクロ経済指標が改定に伴い値が変化することを考慮し、ナウキャスト対象の時点において実際に入手可能であったデータのみを用いた。機械学習手法としては、具体的にはリッジ回帰、LASSO 回帰、エラスティック・ネット、ランダム・フォレストを使用した。

ナウキャストの結果、本稿で用いた機械学習手法は伝統的計量経済手法より精度の高い結果を与えることが明らかになった。加えて、4 つの機械学習手法の中では、ランダム・フォレストが最も精度の高い結果を与えることが示された。また、本稿のナウキャストと公益社団法人日本経済研究センターによるナウキャストの結果とを比較したところ、本稿は 1 次速報値発表の 2 ヶ月前以前のナウキャストにおいて特に優れた結果を与えることが明らかになった。これは、政策担当者が、機械学習手法を用いたナウキャストを利用することで、速報値発表よりもかなり前の時点で経済の現況を反映した意思決定を行える可能性を示す。加えて、本稿のナウキャストと公益社団法人日本経済研究センターのナウキャストとの組み合わせ予測の結果より、本稿によるナウキャストは公益社団法人日本経済研究センターのナウキャストにはない情報を含んでいる可能性も示唆された。さらに、データの状況設定とハイパーパラメータ選択方法の違いが、モデルのパフォーマンス評価に影響を及ぼすことも明らかになった。

同攻会記入欄

書類番号	受付番号
④	

早稲田商学学生懸賞論文 本文

※以下に、表題、本文、図表、注、参考文献を 25 枚以内で記述してください。

機械学習を用いた日本の GDP ナウキャスト

金子 果歩、笹川 俊、福田 和生、宮崎 健太、王思涵

1. イントロダクション

政策担当者は国内総生産(Gross Domestic Product; GDP)や完全失業率、消費者物価指数(Consumer Price Index; CPI)など国の経済状況を表す様々なマクロ経済指標を考慮し経済・金融政策を立案する。これら政策立案に用いられる経済指標の多くは該当期間終了日から起算して、数週間後に発表される。例えば、日本では第 1 四半期(1 月～3 月)GDP の概算値は 5 月 15 日周辺に発表される。一般的に、この時に発表される概算値は「速報値」¹と呼ばれる。そして、四半期終了とその四半期の速報値発表のタイミングの差が、政府や中央銀行による日本経済の現況把握に数週間の遅れをもたらし、それによって経済政策のラグが生じうる。このようなラグを可能な限り解消する為、対象四半期の速報値発表よりも早い時期に精度の高い速報値予測を与えることは、現況にあった経済政策を即時的に立案可能にするという点で重要である。こうした現況予測という考え方は「ナウキャスト」と呼ばれ、将来予測を意味する「フォアキャスト」と区別される(本稿では、以後、ナウキャストとフォアキャストの両方を意味する場合には「予測」という言葉を用いる)。ナウキャストは歴史的には気象の分野で行われてきたが、近年ではマクロ経済の分野でも行われている。

マクロ経済指標のナウキャスト及びその手法の開発、パフォーマンスの評価は世界各国の中央銀行やそこに属する研究者を中心に、活発に行われている(e.g., Matheson, 2010; Fornaro and Luomaranta, 2018; Richardson et al., 2018)。例えば Matheson(2010)は、ダイナミックファ

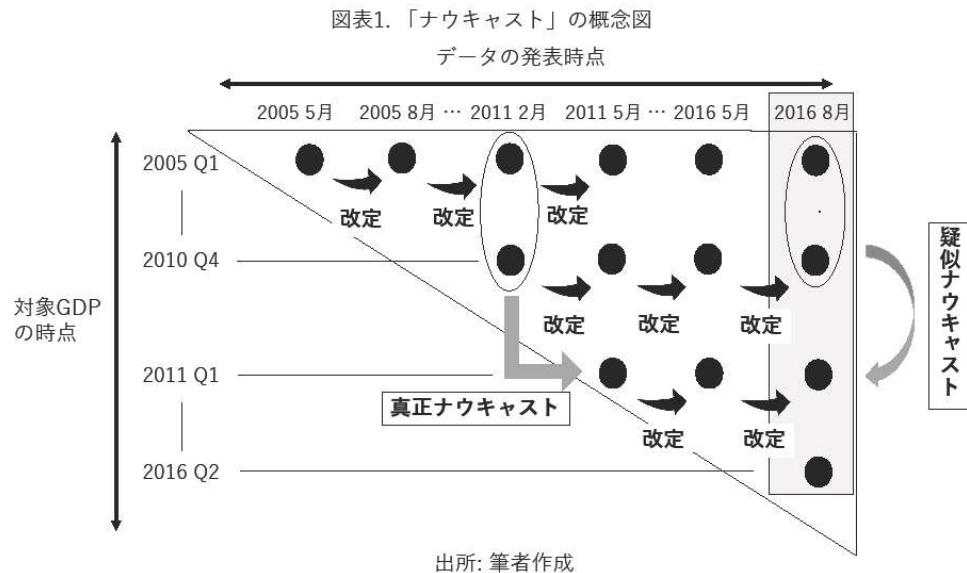
¹ 速報値は指数を作成する為に必要なデータが十分に収集されていない時点で概算された値である。その為、時間の経過とともに、正確な値を算出するためのデータが集まると、速報値の改定がなされていく。速報値に対して、十分にデータが集まり改定が終了した値を「確報値」と呼ぶ。

クターモデル(Dynamic Factor Model; DFM)により、ニュージーランドの実質 GDP 成長率や CPI 上昇率等のナウキャストを行なった。その結果、Matheson(2010)によるモデルは従来のニュージーランド準備銀行による予測と同程度の精度をもたらすことが示された。また、Chikamatsu et al.(2018)は工業生産指数のような月次データを用いた Mixed-Data Sampling 回帰モデルを用いて、日本の GDP の四半期速報値のナウキャストを行なった。その結果、このモデルの予測の精度は標本内平均による予測よりも良いことが明らかになった。

ただ GDP 予測のパフォーマンスに関する多くの研究は、実際に行われる予測の状況設定を想定できていない。多くの研究者は GDP 予測モデルの構築に最新版のデータを用いるが、GDP の速報値は時間の経過とともに改定がなされていくため、実際の予測時点では利用不可能なデータを用いていることになる。例えば Tiffin(2016)では、レバノンにおける 2011 年第 1 四半期の GDP のナウキャストに、2016 年時点まで改定を続けてきた 2011 年の第 1 四半期以前のデータを用いている。図表 1 でいえば、2016 年の 8 月版(灰色部分)から、2011 年第 1 四半期以前の時点のデータ(灰色部分のうち楕円で囲われている箇所)を用いて、2011 年第 1 四半期のナウキャストをしているということである。これは 2011 年第 1 四半期以降に判明した情報を捉え、改定されてきたマクロ経済指数の値を用いて、ナウキャストが行われていることを意味する。これは実務では不可能な方法である。なぜならナウキャストを行う人はその時点でのデータの値は知り得ているが、同じデータの改定後の値をその時点では知り得ないからである。予測のパフォーマンスに関する研究を実際の政策決定に生かすためには、政策立案を行う時点で利用可能なデータのみを用いてナウキャストを行う必要がある。そのような背景の下、近年「真正ナウキャスト」が注目を浴びている。

Bragoli(2017)によると GDP の真正ナウキャストとは、予測者がナウキャストを行う時点で利用可能なデータのみを用いて GDP ナウキャストを行うことである。図表 1 を用いて言えば、2011 年第 1 四半期の GDP のナウキャストを、その発表の直前時期での利用可能なデータである 2011 年 2

月版のデータを用いて行うことである。それに対して、先述した Tiffin(2016)のように、予測を行う時点で知りうるはずのないデータを用いているナウキャストを「疑似ナウキャスト」という。



真正ナウキャストを行っている先行研究として、例えば日本の GDP を対象とした Bragoli(2017)やトルコの GDP を対象とした Modugno et al.(2016)が挙げられる。こうした真正ナウキャストを行っている研究の多くは、手法として DFM を用いている。しかし、Richardson et al.(2018)や Tiffin(2016)が指摘するように、広範な変数から有益な情報を持つ変数を選択し、選択された変数をもとに予測を行うことにおいては機械学習手法が有効である。

実際に、Richardson et al.(2018)では、機械学習手法は自己回帰 (Autoregressive; AR)モデルや DFM などの従来の計量経済手法と比較して良い予測値を与えることが示された。また、イングランドの連邦銀行も同様に機械学習手法を用いてマクロデータの予測を行っている (Chakraborty and Joseph, 2017)。しかし、日本の GDP の真正・疑似ナウキャストに機械学習手法を用いた研究は本稿の知る限り行われていない。

そこで本稿では機械学習手法を用いて、日本の GDP 速報値の真正・疑似ナウキャストを行なった。その結果、先述した先行研究同様、機械学習

手法は伝統的な計量経済手法よりも優れた予測パフォーマンスをもたらすことが示された。さらに、疑似ナウキャストで最も精度の高い手法が真正ナウキャストの最も精度の高い手法と一致するとは限らないことも明らかになった。加えて、ナウキャストのタイミングが GDP 速報値発表の 65 日前よりも早い時期であれば、本稿の手法により、公益社団法人日本経済研究センター(Japan Center for Economic Research; JCER)よりも優れた GDP のナウキャストが可能であることも明らかになった。

本稿の構成は以下の通りである。次節では先行研究のレビュー、3 節では本稿におけるナウキャストのフレームワークの解説、4 節では利用するデータの説明、5 節では用いる手法の紹介を行う。そして 6 節では GDP ナウキャストの結果を与え、7 節ではその結果に基づく結論を述べる。

2. 先行研究

先述したように、マクロ経済指標のナウキャストは、国際金融機関などの専門機関に加えて、数々の研究者によって行われている。ただし、専門機関によるナウキャストは実際の目下の経済状況を対象としているのに対し、研究者によるナウキャストは、モデルの予測精度を評価するために行われており、対象は過去のマクロ経済指標の値である。マクロ経済指標をナウキャストする研究の重要性は認識され始めてきたものの、日本のマクロ経済指標をナウキャストした研究はわずかにしか存在しない。

日本の GDP 成長率をナウキャストした例としては、Hara and Yamane(2013)による静的ファクターモデルを用いた疑似ナウキャスト、Urasawa(2014)による DFM を用いた真正ナウキャスト、Bragoli(2017)による DFM を用いた真正ナウキャストと疑似ナウキャスト、Chikamatsu et al.(2018)による DFM を用いた疑似ナウキャストがある。

ただし、実務で使用するモデルの精度の優劣を議論する際には、疑似ナウキャストは必ずしも適切であるとは限らない。Croushore and Stark(2001)はフォアキャストの文脈で、使用するデータの改定状況によって、モデルの精度の優劣が変化する可能性を指摘している。なぜなら、マ

クロ経済指標の値の改定は、指標の定義の変更やデータソースの変更を含む場合があるため、改定により指標の値が大きく変化するからである。同様の問題がナウキャストの場合にも当てはまる可能性がある。

よって本稿は、使用するデータの改定状況が結果に影響を及ぼす可能性を十分に考慮し、実際に行われるナウキャストの状況設定を完全に模倣した真正ナウキャストを行う。さらには、本稿のモデル構築時点(2019年9月)で手に入った最新のデータを用いた疑似ナウキャストも行い、真正ナウキャストと疑似ナウキャストでは、モデル間の予測精度の優劣が変わるかどうかを検証する。

また、マクロ経済指標の予測においてはモデルの選択も重要である。マクロ経済指標の予測には、回帰予測モデルが用いられることが一般的であった(Hara and Yamane, 2013)。しかし近年では、Giannone et al.(2008)により提案されたDFMが用いられることが多い(Kočenda and Poghosyan, 2018)。なぜなら、DFMは、莫大な数の説明変数の中から、それらの変数を集約的に表す少数の主成分を抽出することで、説明変数間の共線性により生じる問題に対処できるからである。DFMを用いたナウキャストの例としては、Modugno et al.(2016)によるトルコの実質GDP成長率の疑似ナウキャストと真正ナウキャストが挙げられる。この研究は、説明変数にGDPと高い相関を持つことが知られている工業生産指数や輸入数量指数などに加えて、トルコの主要産業である自動車業の状態を表す自動車生産指数や、景況感を問うサーベイからのデータを用いている。その結果、この研究のナウキャストは、国際通貨基金と経済協力開発機構(Organization for Economic Co-operation and Development; OECD)によるナウキャストよりも精度が高いことが明らかになった。

さらに最近では、機械学習がマクロ経済指標の予測に用いられることも増えてきた。Tiffin(2016)によれば、ファクターモデルにおいては、莫大な数の説明変数の候補から、集約的に少数の主成分を抽出する際、欠損値の少ない変数の影響が過度に重視される傾向にある。よって予測因子として抽出された主成分が、実際に最も有効な情報であるとは限らない。一方、機械学習は説明変数間の共線性の問題に、異なる方法で対処する。Tiffin

(2016)は、レバノンの GDP 成長率を、機械学習手法であるエラスティック・ネット(Elastic Net; EN)とランダム・フォレスト(Random Forest; RF)を使用して疑似ナウキャストしており、両者によるナウキャストの精度が十分に高いことを報告している。

この他にも機械学習を用いたマクロ経済指標のナウキャストは幾つも存在する。例えば、Fornaro and Luomaranta(2018)は、フィンランドの GDP 成長率をファクターモデルと機械学習の両方の手法を用いて疑似ナウキャストを行っている。また、Richardson et al.(2018)は、機械学習手法を用いてニュージーランドの GDP 成長率の真正ナウキャストを行っている。その結果、機械学習手法が伝統的計量経済手法と比べて高いパフォーマンスを与えたことが明らかにされている。さらには、様々な機械学習手法による予測値の平均は、それぞれの機械学習手法の精度より高かったことが報告されている。これらの先行研究は、機械学習手法が GDP のナウキャストに非常に有効であることを示唆している。

3. ナウキャストのフレームワーク

本節では、本稿において実際にどのように GDP の疑似・真正ナウキャストを行っていくかの詳細を説明する。まずどちらの種類のナウキャストにも共通する事項として述べなくてはならないのは、本稿のナウキャストは任意の t 期の GDP について、 t 期の GDP に先行して発表されるマクロ経済指標(先行マクロ指標)を用いて行われるということである。これは GDP を従属変数、先行マクロ指標を説明変数とした予測モデルを推定することを意味する。また、 t 期の GDP の予測モデルの推定では $t-1$ 期までの GDP と先行マクロ指標しか用いることができない。最終的な t 期の GDP のナウキャストは、この予測モデルにナウキャストの時点で発表済みの t 期の先行マクロ指標を与えることで行われる。

本稿では、ナウキャストを行う時点ごとに予測モデルの推定を行う。つまり予測モデル 1 個につき 1 個の予測値の算出が行われる。予測モデルの再推定を每期行う理由は、 $t+1$ 期のナウキャストを行う際に、 t 期のナウ

キャストの時点では未発表だった t 期の GDP を予測モデルに反映させることで、精度が向上すると考えられるためである²。

ナウキャストを行う時点は、ナウキャスト対象の GDP 1 次速報値の発表日の 15 日前、45 日前、75 日前に設定する。それぞれの設定で説明変数として利用できる発表済みの先行マクロ指標は異なる。よってそれぞれ 3 つの設定での予測モデルは異なるものになる。

また、注意しなければならないのは、今回利用する GDP と先行マクロ指標には改定が存在し、同じ「 t 期のデータ」であっても、複数の改定された値が存在するということである。そのため、本稿では毎月 1 日時点での各期の変数の値を記録し、月ごとにデータセットを構築する。この月ごとのデータセットをその月の「版」と呼ぶ³。疑似ナウキャストと真正ナウキャストの違いは、端的には利用する変数の版の違いである。

本稿の疑似ナウキャストにおいて、全ての期の全ての変数の値は研究を行う時点(2019 年 9 月版)の値を用いる。一方、真正ナウキャストにおいては、各々のナウキャストの時点において、その時点で利用可能な版の情報のみを用いてナウキャストが行われる。具体的には、 t 期の予測モデルの推定に用いられる「 $t-1$ 期までの説明変数・従属変数」及びその予測モデルに与える「 t 期の説明変数」はそのナウキャストを行う時点における最新の版の値である。

ここで、疑似ナウキャストと真正ナウキャストでは予測対象の「真の値」が異なることも述べておかなければならない。疑似ナウキャストでは t 期のナウキャストにおける真の値は 2019 年 9 月版の t 期の GDP の値になる。したがって精度を評価する際の誤差もこの真の値と t 期の GDP ナウキャストの予測値の差で計算される。これは疑似ナウキャストにおいては 1 つの版の値しか用いておらず、精度評価の際にも版の違いを考慮する必要がないためである。一方、真正ナウキャストでは t 期のナウキャストに

² 各時点での予測モデル推定に用いるデータは 1994 年第 1 四半期からその時点のナウキャストの対象期の 1 期前の四半期までのデータとなる。

³ 1 つの版はスプレッドシート 1 枚のイメージである。

おける予測対象の真の値は t 期の GDP の「1 次速報値」である。これについては真正ナウキャストを行っている Urasawa(2014)などの先行研究に倣ったものである。理由としては、1 次速報値は t 期の予測モデルの推定に用いる版(ナウキャスト時点の最新版⁴)の直後の版の値であるために、それら 2 つの版の間に基準変更が原則存在しないと考えられるためである。

本稿では、疑似・真正ナウキャストを、ベンチマークを含めた後述する手法ごとに 2012 年第 1 四半期から 2019 年第 2 四半期まで、計 30 期分行う。そのうえで予測値の系列について「平均二乗平方根誤差(Root Mean Squared Error; RMSE)」を計算することでその手法を評価する。

$$RMSE = \sqrt{\frac{1}{30} \sum_{t=2012Q1}^{2019Q2} (y_t - \hat{y}_t)^2}$$

ここで \hat{y}_t は t 期の GDP のナウキャストによる予測値であり、 y_t は前述した t 期の GDP の真の値である。

本節の最後に、本稿の目的を改めて述べる。1 つ目は、ナウキャストにおいて機械学習手法がどれほど有効かを明らかにすることである。これは今回の機械学習手法による RMSE と伝統的な計量経済手法による RMSE を比較し、さらに統計的に有意な差が存在するかを検定すること⁵で確認される。2 つ目は、2 つの設定の「ナウキャスト」において複数の機械学習手法のうち、どの手法が最も優れているかを明らかにすることである。これは、機械学習手法間の RMSE の比較で達成される。3 つ目は「疑似ナウキャスト」と「真正ナウキャスト」の間にどれほどの違いが存在するのかを明らかにすることである。これに関してはそれぞれの最も優れた手法に違いがあるのかを確認する。そして 4 つ目は、専門家(JCER)によるナウキャストと機械学習手法のどちらがより優れているを確認することである。これについては 2 つの予測値系列の RMSE を比較する。

⁴ この版は t 期の GDP の 1 次速報値発表の 15 日前、45 日前あるいは 75 日前の版であり当然 t 期の GDP は存在していない。

⁵ Diebold and Mariano(1995)による検定を利用する。

4. データと記述統計

本稿が使用するデータは、OECD の統計データベース OECD. Stat から取得した。以下ではモデルに含めた変数について説明する。

4.1. 変数

本稿のナウキャストの目的となる変数は、四半期毎の実質 GDP 成長率 である。各四半期の GDP は、1 次速報値発表後、何度かの改定を経て確報値となる。3 節で述べた通り、本稿の疑似ナウキャストでは GDP の最新版の改定値を、真正ナウキャストでは 1 次速報値をナウキャストの目的となる変数とする。

本稿は、Stock and Watson(1988)で用いられている複数のマクロ経済指標を参照し、鉱工業生産指数(Index of Industrial Production; IIP)、小売り売上高指数(Index of Retail Trade Volume; RTV)、雇用(Employment; EMP)を説明変数としてモデルに使用した。IIP は国際産業分類における鉱業、製造、電気、ガス、水道の各業種の生産量を 1 つの指数にしたものである。RTV は小売業の総売上高を価格デフレーターで調整した指数である。これらは OECD 基準年を 100 として標準化されている。そして、EMP は新たな求人の人数を測定したものである。IIP 及び EMP は日本の景気と正循環的に変動する傾向があり(Urasawa, 2014)、RTV は GDP の構成要素である家計の消費と大きく連動しうするため、これらは説明変数として妥当と考えた。

次に、物価指数である CPI を含めることでモデルを拡張した。CPI は家計が消費する財とサービスの価格が平均的にどう変化したのかを OECD 基準年を 100 として表したものである。また本稿は GDP と比較的強い相関がある輸出額(Exports; EXP)と輸入額(Imports; IMP)を説明変数として採用した。これらはそれぞれ、時価での輸出入額を表している。さらに、GDP の自己相関を考慮しモデルの精度を向上させるため、説明変数には GDP 成長率の 1 期のラグを加えた(図表 2)。

以上の変数を本稿のナウキャストで用いる。なお、モデルに使用する全て

の変数は対数差分をとる(すなわち目的変数は GDP の四半期成長率となる)。

図表2. 主要な変数の説明

変数				発表の遅延
定義		単位	発表頻度	
鉱工業生産指数	IIP	指数	月毎	約30日
消費者物価指数	CPI	指数	月毎	約30日
輸出額	EXP	千円	月毎	約30日
輸入額	IMP	千円	月毎	約30日
小売り売上高指数	RTV	指数	月毎	約30日
雇用	EMP	千人	月毎	約30日
GDP		円	四半期毎	約45日
GDP成長率	GDP		四半期毎	約45日

注：全ての指標は季節調整された時系列データである。

注：発表の遅延とは、各期間の終了から何日後に実際に速報値が発表されるかを示している。

データ入手元：経済協力開発機構（OECD）

4.2. リアルタイムデータセット

本項では、リアルタイムデータセットについて述べる。データセットは、月毎の最新の変数データをもつ版によって構成される。各版は 1994 年第 1 四半期からその版で利用可能な四半期までの情報を含み、最新の版では 2019 年 9 月における 1994 年第 1 四半期から 2019 年第 2 四半期までの情報が反映されている。

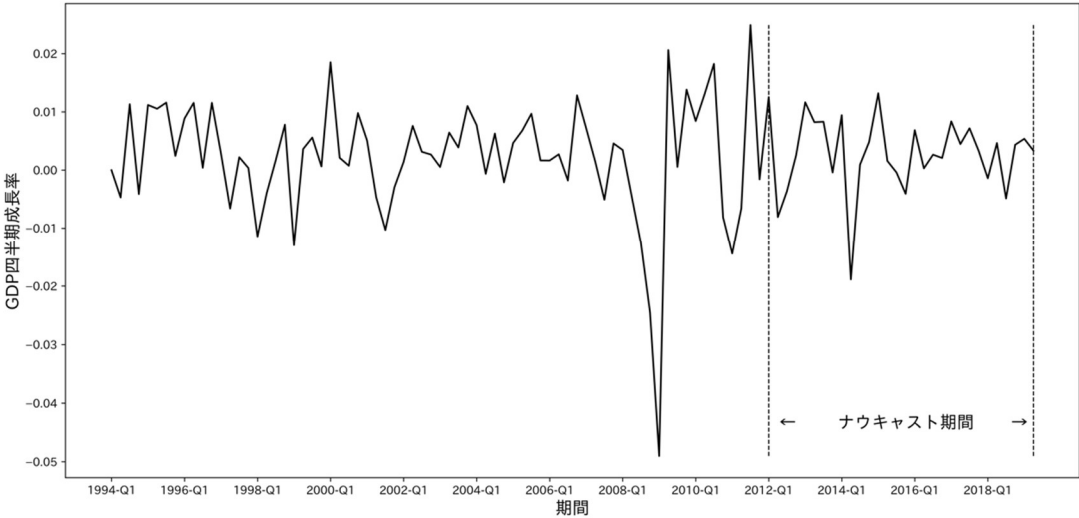
説明変数のうち先述した 6 つの先行マクロ指標は月単位の値である。そのため、本稿では 15 日前ナウキャストの場合、任意の四半期 GDP の対象期間に含まれる第 1 月目、第 2 月目、第 3 月目の 6 つの先行マクロ指標と GDP 成長率の 1 期ラグを説明変数として用いる。45 日前ナウキャストの場合、第 3 月目の先行マクロ指標が未発表であるため、これらは説明変数として用いない。さらに 75 日前ナウキャストの場合は第 2 月目の先行マクロ指標も未発表であるため、これらも説明変数として用いない。

4.3. 記述統計

図表 3 は 2019 年 9 月版の確報値及び最新の改定値に基づく GDP 成長率の推移、図表 4 は本稿で用いた変数の記述統計を示したものである。図表 3

のナウキャスト期間である 2012 年第 1 四半期から 2019 年第 2 四半期においては、2014 年第 2 四半期に成長率の鋭い落ち込みが見られる。これは消費増税を主要な要因にもつ変動と考えられるが、このような変動をうまく捉える情報を取捨選択することができるかどうか、ナウキャストの精度に大きく影響する可能性がある。次節では、実際にその情報の選択を行う各手法について説明していく。

図表3. 2019年9月版に基づくGDP成長率の推移



図表4. 記述統計量

変数		観測数	平均値	標準偏差	最小値	最大値
鉱工業生産指数	IIP	306	103.08	6.15	79.65	119.47
消費者物価指数	CPI	306	98.27	1.57	95.70	102.00
輸出額（百万円）	EXP	306	5,188.89	1,127.93	3,162.80	7,352.60
輸入額（百万円）	IMP	306	4,793.22	1,528.20	2,165.10	7,758.40
小売り売上高指数	RTV	306	100.02	3.76	93.37	119.88
雇用（千人）	EMP	306	64,175,032.70	1,023,468.95	62,630,000.00	67,320,000.00
GDP（百万円）	GDP	102	121,077,578.00	7,524,493.02	105,758,000.00	135,058,000.00

注: モデルの推定の際には、上記の変数をすべて対数変換し、一期前との差分をとったものを用いた。従って、予測の対象はGDPの四半期成長率である。

5. 手法

GDP 成長率の予測において、最小二乗法(Ordinary Least Square; OLS)による多重線形回帰モデルを用いた場合、予測精度が悪くなる可能性がある。その理由は、説明変数となる先行マクロ指標は同じ時期の経済状況を表しているため、モデルに(不完全な)多重共線性の問題が生じるからである。さらに、マクロデータは通常サンプルサイズがそれほど大きくないた

め、係数の最小二乗推定量の分散が非常に大きくなる。その結果、OLS は推定データに含まれないデータの当てはまり(予測問題)に関して優れたものでないことが知られている(Tiffin, 2016)。本稿は、機械学習手法を用いることでこの問題に対処を試みる。

5.1. 予測手法

5.1.1. 罰則付き回帰(正則化回帰)

罰則付き回帰とは、OLS における二乗誤差最小化のプロセスに、係数が過度に大きくなることへの罰則を加えたうえで、それぞれの説明変数の最も適切な係数を推定する手法である。罰則の加え方にはいくつかの種類があり、本稿はその中で、リッジ回帰、ラッソ回帰、エラスティック・ネット(EN)の3つを採用する。それぞれの手法における係数の推定式は次のとおりである。

$$\hat{\beta}_{ridge} = \underset{\beta_0, \dots, \beta_J}{argmin} [\sum_{t=1}^T (y_t - \beta_0 - \sum_{j=1}^J \beta_j x_{tj})^2 + \lambda \sum_{j=1}^J \beta_j^2]$$

$$\hat{\beta}_{LASSO} = \underset{\beta_0, \dots, \beta_J}{argmin} [\sum_{t=1}^T (y_t - \beta_0 - \sum_{j=1}^J \beta_j x_{tj})^2 + \lambda \sum_{j=1}^J |\beta_j|]$$

$$\hat{\beta}_{EN} = \underset{\beta_0, \dots, \beta_J}{argmin} [\sum_{t=1}^T (y_t - \beta_0 - \sum_{j=1}^J \beta_j x_{tj})^2 + \lambda \{(1 - \alpha) \sum_{j=1}^J \beta_j^2 + \alpha \sum_{j=1}^J |\beta_j|\}]$$

またそれぞれの手法における予測値は

$$\hat{y}_t = \hat{\beta}_0 + \sum_{j=1}^J \hat{\beta}_j x_{tj}$$

となる。ここで、 $\hat{\beta}_{ridge}, \hat{\beta}_{LASSO}, \hat{\beta}_{EN}$ はそれぞれの手法で推定される説明変数の係数($\hat{\beta}_0, \dots, \hat{\beta}_J$)からなるベクトルであり、 y_t は推定時点における t 期の実質 GDP 成長率、 x_{t1}, \dots, x_{tJ} は推定時点における t 期の説明変数の値である。また、 λ は「ハイパーパラメータ」と呼ばれるものであり、ここに任意の値を与えることで、係数が大きくなることへの罰則の強さを調節する。一般に、この値を大きくすれば罰則は強くなる。ただ、強い罰則を設定したとしても必ずしも良い予測が実現できるとは限らず、この値はデータに応

じて適切に設定する必要がある。本稿ではこの値を「検証(validation)」により決定する。検証の詳細については次節で説明する。

式で示されているように、リッジ回帰は係数の二乗和、ラッソ回帰は係数の絶対値の和が過大になることを防ぐよう設計されている。これは両者の推定係数に違いをもたらし、ラッソ回帰では予測に重要でない変数について係数が 0 と推定されることが多い。つまり変数選択が行われていることと同じである。また、EN ではこれらの 2 つの手法の罰則項がどちらも利用されており、もう 1 つのハイパーパラメータ α によってその罰則項の影響の割合が決定される。

これらの罰則付き回帰では、係数の不偏性が失われる代わりに、その分散が OLS よりも小さくなる。したがって全体としては推定データに含まれないデータに関して OLS よりもモデルの当てはまりがよくなる可能性がある (Hastie et al., 2009)。また、係数が縮小されるため、説明変数が極端な値をとった場合に予測値の過度な反応を防ぐことができるといったメリットも存在する⁶。

5.1.2. ランダム・フォレスト (RF)

RF とは決定木を応用した手法である⁷。決定木による予測は、「説明変数に閾値を設定し、それを境に説明変数の定義域を 2 つのエリアに分割する」という流れを何度も繰り返し、最終的にその定義域を数個から数十個のエリアに分割したうえで、そのエリアごとに予測値を設定するというアイデアに基づいている。ここで各エリアの予測値は「そのエリアに存在する従属変数の平均値」となる。また、分割の際の説明変数の閾値には、あらゆる値の候補の中で「各エリアの平均二乗誤差の加重和が最も小さくなるもの」が選ばれる。

RF では、標本および利用する変数が無作為に変更されたうえで、決定木

⁶ 罰則付き回帰については Hoerl and Kennard(1970)、Tibshirani(1996)、Zou and Hastie(2005)などが詳しい。

⁷ Breiman(2001)により考案された手法である。

がいくつも構築される⁸。そのうえで、予測値が以下の通り各決定木による予測値の平均で決定される。

$$\hat{y}_t = \frac{1}{B} \sum_{b=1}^B \hat{y}_t^b = \frac{1}{B} \sum_{b=1}^B \text{tree}^b(\mathbf{X}_t)$$

構築される木の本数は B で表され、構築後のそれぞれの決定木に説明変数のベクトル \mathbf{X}_t を入力したときの従属変数の予測値が $\text{tree}^b(\mathbf{X}_t)$ および \hat{y}_t^b で表されている (b はその決定木の番号を示している)。木の本数 B については任意の値を設定することができる。本稿では B を RF のハイパーパラメータとして考え、罰則付き回帰における λ 同様、検証で決定する

また、RF では、予測者は各決定木をどこまで「成長」させるか、つまり各決定木で説明変数の定義域を何回分割させるか(どこまで細かく分解するか)を「木の深さ」として設定する必要がある。本稿はこの木の深さもハイパーパラメータとして考え、検証によって決定する。

RF はデータの非線形関係や相互作用を捉えることができるため、OLS やその線形性を踏襲している罰則付き回帰とは大きく異なるアルゴリズムといえる。そのため、説明変数と従属変数が仮に非線形の関係であった場合、罰則付き回帰よりも優れた精度を与える可能性が高い。また仮に、RF を構成するそれぞれの決定木が標本にオーバーフィッティングしたとしても、それらが最終的に平均化されるために、全体としては予測値の分散を抑えることができるという特徴を持っている(Breiman, 2001)。

5.1.3. 組み合わせ予測

Bates and Granger(1969)などによると、異なる情報を含んだ複数の予測値の相加・加重平均は、元の予測値よりも良い予測値となる傾向がある。これは、複数の予測値系列の相加・加重平均による予測値系列の分散

⁸ 標本および変数の変更が行われる理由としては、決定木に多様性を持たせることで決定木による予測値間の相関を弱め、平均後の予測値の分散を下げるためである。本稿では各決定木において、標本は復元抽出で同数を、変数は全変数のうち 33% をそれぞれ無作為に選択した。

が、元の予測値系列の分散よりも小さくなる可能性があるためである。

本稿では先述した機械学習手法により、複数の予測値系列を得ることになる。各予測値系列は、同じ変数を用いているものの説明変数と従属変数の関係性を異なった形で推定した予測モデルによるものである。したがって、予測値系列が保持している情報に違いがあると考えることができ、これらを平均して新たな予測値系列を得た場合、その精度は元の予測値系列の精度よりも高い可能性がある。よって、本稿では予測値系列の相加平均を計算し、新たな予測値系列を得る。ここで加重平均でなく相加平均を用いる理由は、予測をする前の時点では、どの予測値系列が優れていて、大きな重みをかけるべきかに関して予測者は知りえないためである。

5.2. 検証(ハイパーパラメータの選択)

先述したとおり、今回利用する機械学習手法には、予測者が設定しなければならないハイパーパラメータが存在する。予測精度はこの値によって大きく左右されるため、適切なハイパーパラメータ選択を行うことは手法の選択同様、良い予測をするうえで極めて重要な要素である。

ハイパーパラメータ選択は、一般的に「交差検証法(Cross Validation)」で行われる。これは推定データから一部の「検証データ」を抽出したうえで、残りのデータ(「検証用推定データ」)で様々なハイパーパラメータを用いた複数の予測モデルを推定して検証データの予測を行い、最も誤差の小さくなるハイパーパラメータを採用する方法である。しかしながら今回はデータの順序に意味がある時系列データを扱うため、推定データから無作為に検証データを抽出した場合、データの順序が崩れ適切な検証ができないと考えられる。そこで、本稿は Hyndman and Athanasopoulos(2018)で紹介されている「時系列交差検証」と呼ばれる手法を利用し、この問題を回避する⁹。この検証は、1回の予測ごとに行われ、その都度最適なハイ

⁹ 「繰り返し予測起源による検証」とも言われる。本稿では1回の検証につき検証データを12期分とした。

パーパラメータが選択される。つまり、1つの予測値の算出ごとに1回ハイパーパラメータ選択が行われる。

ここでさらに検討しなければならないのは、真正ナウキャストのそれぞれの検証において、どの版を用いるかである。実務における状況設定を反映した真正ナウキャストにおいて本稿が想定する予測者は、ナウキャストを行う時点までに発表されているすべての版の情報を得ることができる。そこで、本稿では2つの戦略を提案する。

1つ目は「予測時点版検証」である。これは検証で利用するすべてのデータについて、予測時点での最新の版のみを用いる戦略である。この戦略の優れた点としては、変数の基準変更がない点が挙げられる。3節にて説明した通り、実際の予測モデルは予測時点での最新版(予測時点版)を用いて推定される。そのため、検証用推定データと推定データの変数の間に基準変更がないという意味で良いハイパーパラメータ選択を行えると考えられる。

2つ目は「1次速報値データ検証」である。これは検証データの従属変数に1次速報値版を利用し、検証データの説明変数および検証用推定データにはその前時点の版¹⁰を用いる戦略である。この戦略の優れた点としては、検証データの従属変数が予測の目的となる変数と同じ1次速報値になることが挙げられる。さらに検証用推定データはその直前の時点の版のデータであり、真正ナウキャストと完全に同じ設定で検証が行われるといえる。したがって、この戦略では真正ナウキャストの設定を重視したうえでのハイパーパラメータ選択が行うことができると考えられる。

本稿の真正ナウキャストにおいては、4つの機械学習手法について、2つの戦略で検証を行い、それぞれ予測値系列を出す。そして組み合わせ予測についても、戦略ごとに4つの予測値系列の相加平均を算出する。また、2つの戦略・4つの手法の下で求められた計8つの予測値系列の相加平均も算出する。これらの結果を見ることで、どの機械学習手法が優れたナウキ

¹⁰ 15日前予測の場合は15日前の版を検証用推定データに用いる(45日前予測、75日前予測だとそれぞれ45日前、75日前の版)。

キャストを行うかに加え、どちらの検証戦略が優れているのかについても知見を得ることができる。

6. 結果

図表 5 は速報値発表の 15 日前、45 日前、75 日前の疑似ナウキャストと真正ナウキャストの RMSE を示す。ベンチマークモデルとして、伝統的計量経済手法である 1 期の GDP 成長率のラグを説明変数とした AR(1)モデルとそこに先行マクロ指標を加えた ARX モデルを用いる。まず、15 日前、45 日前、75 日前の疑似ナウキャストと真正ナウキャストすべてにおいて、いずれかの機械学習手法が 2 つのベンチマークモデルよりも低い RMSE を与えた。特に 15 日前の真正ナウキャストでは、多くの機械学習モデルが Diebold-Mariano 検定(Diebold and Mariano, 1995)により AR(1)と統計的に有意に異なる予測精度を与えることが分かった。この結果は、機械学習手法を用いることで伝統的な計量経済学手法よりも優れた GDP のナウキャストができる可能性を示唆している。

次に、15 日前の結果について、疑似ナウキャストでは RMSE が 0.00536 である EN が最も精度の高いモデルであった。一方で、真正ナウキャストでは RMSE が 0.00475 である予測時点版検証を用いた RF が最も精度の高いモデルであり、疑似ナウキャストとは異なるモデルが最良なモデルであった。同様に 45 日前の疑似ナウキャストでの RMSE が最小のモデルがラッソ回帰であったのに対し、真正ナウキャストでの RMSE が最小のモデルは 1 次速報値データ検証を用いた RF であった。この結果より、疑似ナウキャストで最良の結果をもたらすモデルは必ずしも真正ナウキャストにおいて最良の結果をもたらすとは限らないことがいえる。つまり、これまでの多くの研究では疑似ナウキャストにおけるモデルのパフォーマンスのみを議論しているが、それだけでは実務における優れたナウキャストに関する知見を与えることにはならない可能性がある。

組み合わせ予測の結果は、疑似ナウキャストにおいて 15 日前と 45 日前で 4 つの単独の機械学習モデルよりも優れたナウキャストを与えたことを

図表5. 平均二乗平方根誤差

		15日前			45日前			75日前		
		真正ナウキャスト			疑似ナウキャスト			真正ナウキャスト		
		真正ナウキャスト	疑似ナウキャスト	真正ナウキャスト	疑似ナウキャスト	真正ナウキャスト	疑似ナウキャスト	真正ナウキャスト	疑似ナウキャスト	真正ナウキャスト
AR(1)		0.00672		0.00661	0.00672	0.00662	0.00672	0.00672	0.00672	0.00652
ARX		0.00614		0.00642	0.00581	0.00623	0.00651	0.00651	0.00651	0.01699
LASSO	- 予測時点版検証	0.00546		0.00523 *	0.00557	0.00570	0.00676	0.00676	0.00676	0.01394
	- 1次速報値 データ検証	-		0.00546	-	0.00634	-	-	-	0.01379
Ridge	- 予測時点版検証	0.00538		0.00531 *	0.00562	0.00587	0.00679	0.00582 **	0.00582 **	0.00582 **
	- 1次速報値 データ検証	-		0.00487	-	0.00568	-	-	-	0.00568
EN	- 予測時点版検証	0.00536		0.00517	0.00564	0.00607	0.00696	0.00662	0.00662	0.00662
	- 1次速報値 データ検証	-		0.00499	-	0.00597	-	-	-	0.00646
RF	- 予測時点版検証	0.00544		0.00475 *	0.00582	0.00560	0.00627	0.00561	0.00561	0.00561
	- 1次速報値 データ検証	-		0.00488 **	-	0.00555	-	-	-	0.00517
AllComb										
PComb		0.00527		0.00484 *	0.00555	0.00573	0.00657	0.00665	0.00657	0.00665
FComb		-		0.00483 *	-	0.00581	-	0.00661	-	0.00661
		-		0.00489	-	0.00570	-	0.00676	-	0.00676

注: 各ナウキャストの設定においてRMSEが最小であるものを太字で示した。AllCombは予測時点版検証・1次速報値データ検証を用いたLASSO、Ridge、EN、RFの計8つの手法の組み合わせ予測のRMSEである。PCombは予測時点版検証を用いた4つの機械学習手法、FCombは予測時点版検証を用いた4つの機械学習手法の組み合わせ予測のRMSEである。各モデルの予測精度がAR(1)と有意に異なるかどうかを検定するDiebold-Mariano testを行い、10%統計的に有意であるものに*を、5%統計的に有意であるものに**を記した。

示した。真正ナウキャストにおいては3種類の組み合わせ予測を求めているが、15日前、45日前、75日前のいずれにおいても機械学習手法で最も優れたモデルのRMSEを下回ることにはなかった。この結果からも疑似ナウキャストと真正ナウキャストに違いがあることがわかる。

図表6. JCERのナウキャスト結果との比較

	15日前	45日前	65日前	75日前	80日前
OurBest	0.00475	0.00555	0.00517	0.00517	-
JCER	0.00334	0.00481	0.00481	0.00554	0.00554

注: OurBestでは図表5で示した本稿における真正ナウキャストで精度が最も優れたモデルのRMSEを示しており、15日前は予測時点版検証を用いたRF、45日前と75日前は1次速報値データ検証を用いたRFの値である。なお、本稿の65日前には75日前の値を代入した。JCERでは日本経済研究センターのGDPナウキャストのRMSEを示した。なお、JCERの45日前には65日前の値を、75日前には80日前の値を代入した。

出所: <https://www.jcer.or.jp/esp/backno.html>
<https://www.jcer.or.jp/economic-forecast/short-term>

図表7. JCERのナウキャストとの組み合わせ予測

	JCER15日前と本稿15日前	JCER65日前と本稿75日前	JCER80日前と本稿75日前
All	0.00460	0.00619	0.00628
P	0.00446	0.00597	0.00614
F	0.00442	0.00579	0.00595
LASSO-P	0.00409	0.00768	0.00808
Ridge-P	0.00411	0.00497	0.00537
EN-P	0.00405	0.00510	0.00552
RF-P	0.00369	0.00486	0.00519
LASSO-F	0.00419	0.00751	0.00790
Ridge-F	0.00392	0.00480	0.00515
EN-F	0.00397	0.00495	0.00533
RF-F	0.00374	0.00463	0.00500

注: JCERのナウキャストと本稿の真正ナウキャストの組み合わせ予測によるRMSEを示した。図表6に示したJCERよりも小さい組み合わせ予測のRMSEを太字で示した。Allは本稿の真正ナウキャストのうち予測時点版検証・1次速報値データ検証を用いたLASSO、Ridge、EN、RFの計8つの手法とJCERの組み合わせ予測、Pは本稿の真正ナウキャストの予測時点版検証のみの計4つの手法とJCERの組み合わせ予測、Fは本稿の真正ナウキャストの1次速報値データ検証のみの計4つの手法とJCERの組み合わせ予測である。本稿の各ナウキャスト手法とJCERの組み合わせ予測も個別に算出した(-Pは予測時点版検証、-Fは1次速報値データ検証を用いたものを表す)。

次に、JCERと本稿のナウキャストを比較する(図表6)。JCERは毎期GDPの速報値発表の15日前と65日前と80日前にナウキャストを行っている。このときJCERは当時利用可能な最新データのみを用いているためJCERのナウキャストは真正ナウキャストであるといえる。まずJCERの15日前のナウキャストと本稿の15日前の真正ナウキャストで比較すると、JCERのRMSEはどの機械学習手法よりも小さい結果となった。

またJCERの65日前のナウキャストと本稿の75日前の真正ナウキャストで比較しても、JCERの方が優れた精度のナウキャストをしている結果となった。しかしJCERの80日前のナウキャストと本稿の75日前の真正

ナウキャストで比較すると、本稿の 1 次速報値データ検証をした RF の RMSE が 0.00517 であり、JCER の RMSE である 0.00554 を下回った。つまり、GDP 速報値発表の 65 日前よりも早い時期であれば、本稿で提案する手法は JCER よりも優れたナウキャストを与えるといえる。JCER の 15 日前と 65 日前のナウキャストが本稿のナウキャストより精度が高い理由として、GDP 速報値発表時期に近づくほど JCER が本稿の入手できていない情報を多く用いてことが考えられる。

さらに、JCER の予測値系列と本稿の真正ナウキャストの予測値系列の組み合わせ予測を計算することにより、本稿のナウキャストが JCER の結果を改善するだけの情報を有しているかを確認する。本稿におけるすべての機械学習モデルの予測値系列、予測時点版検証を用いた機械学習モデルの予測値系列、1 次速報値データ検証を用いた機械学習モデルの予測値系列、それぞれ単独の機械学習モデルの予測値系列と JCER の予測値系列の平均を求め計 10 種類の組み合わせ予測を計算した。15 日前以外の本稿のナウキャストのタイミングと JCER のナウキャストのタイミングは異なるため、JCER の 65 日前と 80 日前については本稿の 75 日前の予測値系列との組み合わせ予測を計算した。図表 7 はそれらの RMSE を示している。

まず 15 日前ナウキャストとの組み合わせ予測において、いずれも元の JCER のナウキャストのパフォーマンスを改善することはなかった。次に JCER の 65 日前ナウキャストとの組み合わせ予測に注目すると、本稿の 75 日前の 1 次速報値データ検証を用いたリッジ回帰との組み合わせ予測と RF との組み合わせ予測が元の JCER の 65 日前ナウキャストの精度を上回った。これは、本稿が JCER に比べて 10 日前のデータを用いたにもかかわらず、本稿の予測値系列に JCER にはない情報が含まれているためと考えられる。JCER の 80 日前ナウキャストとの組み合わせ予測では、予測時点版検証を用いたリッジ回帰、EN、RF、1 次速報値データ検証を用いたリッジ回帰、EN、RF と JCER とのそれぞれの組み合わせ予測が元の JCER の 80 日前ナウキャストの精度を上回った。以上の結果より、本稿のナウキャストは JCER の結果を改善しうる情報を有していることが確認でき、本稿の手法を活用することで実務における GDP ナウキャストのパフォーマ

ンスを向上させる可能性があるといえる¹¹。

7. 結論

マクロ経済指標は、該当期間終了から数週間後にその 1 次速報値が発表される。しかし、政策担当者は目下の経済状況を反映した政策を速やかに立案する必要があるため、経済の現況を予測するナウキャストが重要であることが認識されている。本稿は、日本の経済をナウキャストした研究には用いられた例のない機械学習を用いて、日本の GDP 成長率の疑似ナウキャスト及び真正ナウキャストを行った。機械学習手法としては、具体的にはリッジ回帰、LASSO 回帰、エラスティック・ネット、ランダム・フォレストを使用した。本稿により得た知見は以下である。

- (1) 速報値発表の 15 日前、45 日前、75 日前における疑似ナウキャストと真正ナウキャストの全てにおいて、いずれかの機械学習手法が伝統的な計量経済手法より優れた結果を与えた。
- (2) 速報値発表の 15 日前の疑似ナウキャストと真正ナウキャストでは、最も優れたパフォーマンスを発揮したモデルが異なったことから、疑似ナウキャストで最良の結果をもたらすモデルが、真正ナウキャストにおいても最良の結果をもたらすとは限らないことが分かった。
- (3) 真正ナウキャストにおいて 2 通りの戦略でハイパーパラメータの選択を行ったところ、同じモデルであっても選択方法によりナウキャストの精度が異なることが分かった。
- (4) 本稿の機械学習手法は、JCER と比較して、速報値発表よりも 2 ヶ月以上前において優れたナウキャスト結果を与えた。
- (5) JCER の 65 日前および 80 日前ナウキャストと本稿の 75 日前のナウキャストを組み合わせたところ、複数の組み合わせ予測の精度は元の JCER のナウキャストの精度を上回った。よって本稿の予測値系列には

¹¹ なお、本稿において考案し、実際に用いたソースコードを <https://github.com/hajimetrics/gdp-nowcasting> に掲載した。

JCER の予測値系列にはない情報が含まれている可能性が示唆された。

一方で、本稿の限界として、まず、説明変数にソフトデータを含めていない点がある。先行研究の多くはナウキャストにおけるソフトデータの有効性を指摘している。次に、本稿は機械学習手法とファクターモデルとの比較は行っていない。機械学習はファクターモデルの問題点に新たな対処法を与えるアプローチであり、ファクターモデルとの比較をした場合機械学習の有効性をさらに議論できた可能性がある。最後に、本稿で用いた手法の他にも、先行研究によりナウキャストに有効であることが示されている機械学習手法が存在する。それらの手法を用いることで、機械学習の有効性をさらに示すことができた可能性がある。

日本の GDP をナウキャストした研究はごくわずかであり、そのほとんどは実務におけるナウキャストを想定していない。本稿は、実務の状況設定を忠実に再現し、日本の GDP のナウキャストにおける機械学習の有用性を示している。これは、政策担当者が、機械学習手法を用いたナウキャストを利用することで、速報値発表よりもかなり前の時点で経済の現況を反映した意思決定を行える可能性を示す。よって本稿は日本の経済にとって非常に重要な発見を提示しているといえる。本稿が経済政策のラグの問題を改善させる一助となること、そして日本の GDP ナウキャストに関する研究に新たな扉を開くことを期待する。

参考文献

Adam Richardson, Thomas van Florenstein Mulder and Tuğrul Vehbi (2018) “Nowcasting New Zealand GDP Using Machine Learning Algorithms,” *Centre for Applied Macroeconomic Analysis Working Paper*, No.47, September 2018

Andrew J. Tiffin (2016) “Seeing in the Dark; A Machine-Learning Approach to Nowcasting in Lebanon,” *IMF Working Paper*, WP/16/56, International Monetary Fund

Arthur E. Hoerl and Robert W. Kennard (1970) “Ridge Regression: Biased

- Estimation for Nonorthogonal Problems,” *Technometrics*, Volume 12, No. 1, pp.55-67
- Chiranjit Chakraborty and Andreas Joseph (2017) “Machine Learning at Central Banks,” *Bank of England Working Paper*, No.674, September 2017
- Daniela Bragoli (2017) “Now-casting the Japanese Economy,” *International Journal of Forecasting*, Volume 33, Issue 2, pp.390-402
- Dean Croushore and Tom Stark (1999), “A Real-Time Data Set for Macroeconomics,” *Federal Reserve Bank of Philadelphia Working Paper*, No. 99-4
- Domenico Giannone, Lucrezia Reichlin and David Small (2008) “Nowcasting: The Real-time Information Content of Macroeconomic Data,” *Journal of Monetary Economics*, Volume 55, pp.665-676
- Evzen Kočenda and Karen Pogosyan (2018) “Nowcasting Real GDP Growth with Business Tendency Surveys Data: A Cross Country Analysis,” *KIER Working Papers*, No.1002, pp.1-26
- Francis X. Diebold and Roberto S. Mariano (1995) “Comparing Predictive Accuracy,” *Journal of Business & Economic Statistics*, Volume 13, Issue 3, pp.253-263
- Hui Zou and Trevor Hastie (2005) “Regularization and Variable Selection via the Elastic Net,” *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, Vol. 67, No. 2, pp. 301-320
- James H. Stock and Mark W. Watson (1988), “A Probability Model of the Coincident Economic Indicators,” in *Leading Economic Indicators: New Approaches and Forecasting Records*, eds. K.Lahiri and G. Moore, Cambridge: Cambridge University Press, pp.63-89
- James M. Bates and Clive W. J. Granger (1969) “The Combination of Forecasts,” *Operational Research Quarterly*, Vol 20, pp.451-468
- Kyosuke Chikamatsu, Naohisa Hirakata, Yosuke Kido and Kazuki Otaka

- (2018) “Nowcasting Japanese GDPs,” *Bank of Japan Working Paper Series*, No. 18-E-18, November 2018
- Leo Breiman (2001) "Random Forests," *Machine Learning*, Vol 45, Issue 1, pp.5-32
- Michele Modugno, Barış SoybilgenbEge and Ege Yazgan (2016) “Nowcasting Turkish GDP and News Decomposition,” *International Journal of Forecasting*, Volume 32, Issue 4, pp.1369-1384
- Naoko Hara and Shotaro Yamane (2013) “New Monthly Estimation Approach for Nowcasting GDP Growth: The Case of Japan,” *Bank of Japan Working Paper Series*, No. 13-E-14, October 2013
- Paolo Fornaro and Henri Luomaranta (2018) “Nowcasting Finnish Real Economic Activity: A Machine Learning Approach,” *Empirical Economics*, Volume 30, pp.1-17
- Robert Tibshirani (1996) "Regression Shrinkage and Selection via the Lasso," *Journal of the Royal Statistical Society: Series B (Methodological)*, Vol. 58, No. 1, pp. 267-288
- Rob J Hyndman and George Athanasopoulos (2018) *Forecasting: Principles and Practice Second edition*, Melbourne: OTexts
- Satoshi Urasawa (2014), “Real-time GDP Forecasting for Japan: A Dynamic Factor Model Approach,” *Journal of the Japanese and International Economies*, Volume 34, pp.116-134
- Trevor Hastie, Robert Tibshirani, and Jerome Friedman (2009) *The Elements of Statistical Learning: Data Mining, Inference, and Prediction Second Edition*, New York; Springer (杉山 将・井手 剛・神 島 敏弘・栗田 多喜夫・前田 英作 監訳「統計的学習の基礎 —データマイニング・推論・予測—」 共立出版、2014)
- Troy D. Matheson (2010) “An Analysis of the Informational Content of New Zealand Data Releases: The Importance of Business Opinion Surveys,” *Economic Modelling*, Volume 27, Issue 1, pp.304-314

付録. GDPナウキャスト値

