

景気単語極性辞書の構築とその応用

五島 圭一[†]・新谷 元嗣[†]・高村 大也^{††}

本稿では、マクロ経済分野に特化した景気単語極性辞書の構築とその応用例を示す。日本経済新聞の新聞記事コーパスから景気関連の単語の選定し、複数人のエコノミストによるクラスラベルの付与を行った。さらに教師あり学習で収録単語を補完した。また、構築した極性辞書の応用例として、景気動向を迅速に捕捉することが可能な経済ニュース指数を作成した。

キーワード：極性辞書、マクロ経済学、景気循環

Sentiment Dictionary for Business Cycle Analysis and its Applications

KEIICHI GOSHIMA[†], MOTOTSUGU SHINTANI[†] and HIROYA TAKAMURA^{††}

In this study, we construct a sentiment dictionary for the macroeconomic domain and present its applications. Our dictionary contains words selected by several economists from a corpus of newspaper articles on topics related to the economy. This was supplemented with additional words by using supervised learning. We use our sentiment dictionary to construct a daily business cycle index designed to capture the current state of the economy in a timely manner.

Key Words: *Sentiment Dictionary, Macroeconomics, Business Cycle*

1 はじめに

適切な経済政策の運営には、足もとの景気動向をできるだけ迅速かつ正確に把握することが肝要である。ところが、一国のマクロ経済活動の最も重要な指標である GDP（国内総生産）は、3 か月に 1 度しか観測されない四半期データであり、対象とする四半期の終了から一次速報値の公表までに約 6 週間を要する。景気の現状判断のための速報性を重視した景気動向指数の一致指数については、より観測頻度の高い月次データである鉱工業生産指数、有効求人倍率、全産業の営業利益、小売業や卸売業の商業販売額等、合計 10 個の一致系列を統合することで計算されるが、それでも対象とする月から 1 か月以上公表が遅れてしまう。そこで景気動向指数の採用系列のような数値による伝統的な構造化データに代わり、テキスト情報、位置情報、決済情

[†] 東京大学エコノミックコンサルティング株式会社, UTokyo Economic Consulting Inc.

^{††} 産業技術総合研究所, National Institute of Advanced Industrial Science and Technology

報, 衛星画像情報を含む非構造化データやオルタナティブデータを利用することによって高頻度かつ高精度に足もとの経済活動を把握する, いわゆるナウキャストの試みが近年進められている. 特に, 直近の新型コロナウイルス感染症拡大に伴う景気後退局面では, 位置情報のデータを用いた生産活動のナウキャストやクレジットカード決済情報のデータを用いた消費活動のナウキャストが注目を集めた(王 他 2021; 日本銀行 2020). 本稿では, テキスト情報のデータを用いた景気動向のナウキャストや将来予測に有用であると考えられるマクロ経済分野の新しい極性辞書(以下, 景気単語極性辞書)の構築とその応用可能性に関する技術報告を行う.

伝統的なデータでは直接観測することが困難であったマクロ経済の変動要因を, 政府や民間のテキストデータから定量化する研究は, 当初は英語圏の欧米経済の分析を中心に進められてきた. 著名な研究例としては, 新聞のニュース記事中の経済政策の不確実性に関する単語の出現頻度を数値化した経済政策の不確実性指数があり, この指数はマクロ経済活動に占める不確実性の波及効果の分析で頻繁に利用されている(Baker et al. 2016). また, 中央銀行の議事録に対してトピックモデルを適用し, 様々なマクロ経済現象の中から, 特に中央銀行が重視するトピックを抽出することによって金融政策の要因分析を行う試みもある(Hansen and McMahon 2016). さらにGDP, インフレ率, 失業率等の主要なマクロ経済変数の将来予測についても, 伝統的なデータのみを用いる従来の方法に比べて, 新聞記事等のテキストデータを活用することで予測力が改善されるという共通認識も広まっている(Kalamara et al. 2022). これらのマクロ経済分析で用いられたテキストデータの数値化では, 計算言語学や自然言語処理等の分野で開発されてきた分析手法や言語資源等が重要な役割を担ってきた.

このようにテキストデータを利用した分析は近年のマクロ経済研究に大きく寄与している一方で, 多くの研究は英語を対象としたものであり, 日本語を対象とした研究の数は相対的に少ない. この理由としては, 日本語で記述されたテキストデータの分析を研究対象とするマクロ経済学者が少ないことに加えて, マクロ経済分野に適した日本語の言語資源の不足がある. 例えば, 前述の(Kalamara et al. 2022)は, 英語の新聞ニュース記事を用いた英国マクロ経済の将来予測の分析であるが, 複数の極性辞書やルールからテキストデータを指数化することによって予測変数を作成している. このように, 英語の分析では評価の定まった辞書やルールが利用できる一方で, マクロ経済分野で確立された日本語の極性辞書は存在しない¹. さらに, 英語の先行研究と同じ一般的な手続きを採用しようとしても, 経済分析に適した単語分割やストップワードの除去等の日本語の前処理方法に関する研究の蓄積も十分とはいえない. このような日本語の言語資源の不足は, テキストデータを用いたマクロ経済研究の敷居を高める要因となっ

¹ マクロ経済学と隣接したドメインであるファイナンス(金融)分野では(五島, 高橋 2017)や(Ito et al. 2018)が極性辞書を構築している. また, 同分野の英語の極性辞書としては(Loughran and McDonald 2011)が広く用いられている. マクロ経済学は一国全体の財(モノ)やサービスの生産と消費の構造や貨幣の流れを解明し, 国民厚生を最大化を目的としたマクロ経済政策を研究する学問分野であるのに対し, ファイナンスは資産価格決定や投資家のポートフォリオ選択, 企業の資金調達等のミクロ行動を主に分析対象とする学問分野である.

ている。以上の問題意識から、マクロ経済分野、特に景気動向の現状把握と将来予測に適した日本語の景気単語極性辞書の構築は、テキストデータを用いた経済研究の発展に貢献できると考えられる²。

2 関連研究

日本語の極性辞書の構築を試みた先行研究は数多くあり、その一部では、研究成果として極性辞書が公開されている (小林 他 2005; Takamura et al. 2005; Kanayama and Nasukawa 2006; Kaji and Kitsuregawa 2007)。ただし、こうした先行研究で公開されている極性辞書は一般的な用途を想定したものが多く、特定のドメインに対してはうまく機能しない可能性がある。実際に、テキスト分析におけるドメイン考慮の重要性は周知であり、(Turney 2002) や (Qiu et al. 2009) では一般的な辞書で全てのドメインに対応することは難しく、ドメインごとに辞書を新たに構築する必要があると論じている。また、(Liu 2012) では、センチメント分析はドメイン依存性が高いことを指摘している。以上の観点から、本稿のようにマクロ経済ドメインに特化した極性辞書の構築は重要な課題であると言える。

3 景気単語極性辞書に収録する単語候補の選択

本章では、景気単語極性辞書に収録する単語の選択手順について記述する。本稿では、マクロ経済分析、特に景気循環の分析に適した極性辞書の構築を目的としていることから、経済や景気と関連のある単語を取得するために株式会社日本経済新聞社が発行する日本経済新聞の記事データ (朝刊・夕刊・電子版) を利用する。使用する新聞記事データの期間は 1981 年 10 月から 2020 年 12 月であり、約 510 万記事が収録されている。また、新聞記事データに含まれる単語数については、異なり数が 1,631,065 語、延べ数が 1,621,203,196 語である。はじめに、「景気」、「経済」もしくはそれらの合成語と各単語の共起回数を算出する。具体的には、各単語について、「景気」、「経済」もしくはそれらの合成語の少なくともいずれか一つと共起した記事数を共起回数とする。さらに、共起回数による単語のランキングを作成する。ここで、考慮する単語は名詞、動詞、形容詞のみとし、またすべての単語は基本形に変換されている。単語分割、品詞推定および基本形の推定には MeCab を利用した。また、マクロ経済に関連する新語や複合名詞を収録することが重要であるため、形態素解析辞書には mecab-ipadic-NEologd を採用した (佐藤 他 2017)。次に、共起回数のランキング上位から、経済や景気と関連があり、収録の候補

² 本稿で構築した景気単語極性辞書は株式会社日本経済新聞社のウェブページ (<https://nkbb.nikkei.co.jp/alternative/service/dictionary/>) から提供されている。

となる単語を、マクロ経済学及びファイナンスをそれぞれの専門分野とする筆者2人が独立に2,000語を選択した。ただし、過去の特定のイベントと関連する単語を収録してしまうと、過去に過剰適合した経済分析になってしまうことから、「リーマンショック」や「新型コロナウイルス」などの特定のイベント固有の単語は使用せず、景気に関する一般的な単語のみを選択することとした。それぞれが選択した単語の重複を排除した後、調整を行い、総単語数が3,000語となるように、候補単語のリストを作成した。この結果、単語リストの品詞の内訳は形容詞：309語、動詞：989語、名詞：1,702語となった。

4 アノテーションとその信頼性評価

3章で得られた単語リストに収録されている3,000語に対して、景気動向の観点で各単語から連想される意味について、ポジティブ・ニュートラル・ネガティブの3つのクラスラベルの付与を行うアノテーションを行った。本稿のアノテーションには経済システムや景気動向に関する深い理解が必要であることから、アノテータとしてマクロ経済分析に従事する3人の専門家に依頼した。なお景気の見方については、個人の経歴や立場によって異なる可能性があるため、クラスラベルが特定の業種の意見を反映しないように、アノテータはそれぞれ別の職種の専門家（政府・中央銀行エコノミスト、民間金融機関エコノミスト、報道機関エコノミスト）を選定した。アノテータへの指示書は以下のとおりである。

アノテータへの指示書

- 景気動向の観点から、各単語から連想される意味について、ポジティブ/ネガティブ/ニュートラルの3つに分類してください。例：「悪化」→ネガティブ、「回復」→ポジティブ
- 一般的には良い意味でも、景気の文脈では悪い意味の場合はネガティブ、どちらとも言えない場合はニュートラルとしてください。悪い意味についても同様をお願いします。
- 一般的に文脈に依存する単語や係り受けによって意味が反転する単語の場合でも、可能な限り、景気が良い時期や内容に利用される場合が多い単語はポジティブ、悪い時期や内容に利用される場合が多い単語はネガティブとしてください。どちらとも言えない単語については、ニュートラルとしてください。例：「不確実」は一般的には係り受けによってどちらの意味にもなりうるが、景気の文脈だと悪い局面に使用される場合が多いのでネガティブと判断する。
- ポジティブ/ネガティブとニュートラルの境目にあると考えられる場合は、ニュートラルに分類してください。
- ひらがなについては、複数の漢字が当てはまる場合であっても、景気の文脈で最も多く使われる使用法を念頭に選んでください。

表1には、各アノテータによるクラス分類の結果とそれらを多数決によって集計した結果が示されている。アノテータ間で各クラスへの分類割合は多少変わるものの、ニュートラルが一番高く、次いでネガティブ、ポジティブの順に低くなる傾向が共通に観察される。分類割合が異なっている要因の一つとしては、各アノテータの判断基準の閾値の違いが想定される。

また、表2にはアノテータ間の評価に関する混同行列、表3には評価の一致度が示されている。表2からは、クラス分類がアノテータ間で不一致だった単語のほとんどがポジティブとニュートラルあるいはネガティブとニュートラルの間であり、ポジティブとネガティブの間で不一致だった単語はわずかであることがわかる。実際に逆のクラスラベルを付与されたのは、「カネ

アノテータ	ポジティブ	ニュートラル	ネガティブ
A	372	2,203	425
B	191	2,496	313
C	383	2,110	507
多数決	282	2,317	401

表1 アノテーションの集計結果

		B		
		ネガティブ	ニュートラル	ポジティブ
A	ネガティブ	242	183	0
	ニュートラル	70	2,087	46
	ポジティブ	1	226	145

(a) A vs. B

		C		
		ネガティブ	ニュートラル	ポジティブ
A	ネガティブ	329	90	6
	ニュートラル	175	1,885	143
	ポジティブ	3	135	234

(b) A vs. C

		C		
		ネガティブ	ニュートラル	ポジティブ
B	ネガティブ	248	59	6
	ニュートラル	258	2,008	230
	ポジティブ	1	43	147

(c) B vs. C

表2 アノテータ間の混同行列

	A vs. B	A vs. C	B vs. C	A vs. B vs. C
κ 係数	0.52	0.59	0.49	0.53

表 3 アノテータ間の一致度.

注：各組み合わせの評価の一致度はコーエンの κ 係数，3 人の評価の一致度はフライスの κ 係数により計算されている。

余り」や「過熱感」，「バブル」といった単語であり，これらは立場によって判断が分かれる可能性が高い単語であると考えられる．表 3 によると，評価の一致度を測る κ 係数は 0.5 から 0.6 程度であり，アノテータの立場の違いを考慮に入れた場合には十分に信頼できる水準といえる．これらの結果を踏まえて，アノテータの多数決によって 3,000 語に対するクラスラベルを決定した．

5 収録単語の補完

本章では，4 章で得られた極性辞書への収録単語の補完を行う．現段階では，新聞記事データにおいて「景気」及び「経済」とそれらの合成語と共起した単語のみが辞書に収録されているが，これら以外の単語について取りこぼしの可能性を考慮して辞書を補完することが望ましいだろう．しかしながら，人手によって全ての追加単語を選定するのは多くの時間と労力を要する．このため本稿では単語の埋め込みベクトルと教師あり学習を導入することで単語の補完を行う (Tang et al. 2014a, 2014b; 佐藤 他 2016)．具体的な手順は以下のとおりである．

- (1) はじめに，単語の埋め込みベクトルを作成する．単語ベクトルを作成する際のコーパスとしては，日本経済新聞の新聞記事データを利用している．なお，本稿では極性辞書に基本形のみを収録するため，入力テキストを形態素解析によって基本形に変換している．単語ベクトルを作成する際の手法として，Skip-gram (word2vec)，GloVe，fastText の 3 種類を採用した (Mikolov et al. 2013; Pennington et al. 2014; Bojanowski et al. 2017)．ハイパーパラメータはデフォルトを指定しているが，ベクトルの次元は 100, 200, 400, 800 の 4 種類を試行している．
- (2) 次に，単語リストの 3,000 語の単語ベクトルを入力，クラスラベルを出力として教師あり学習を行う．教師あり学習の手法としては，多項ロジスティック回帰 (MLR)，サポートベクトルマシン (SVM)，順伝播型ニューラルネットワーク (FFNN)，LightGBM を用いる．なお，SVM に関しては線形カーネル，ガウシアンカーネル，多項式カーネル，シグモイドカーネルの 4 つのカーネル関数を試行するとともに，one-versus-rest 法によっ

て3クラス分類を行った³。また、表1が示すとおりクラスラベルの割合にばらつきがあることから、その逆数を掛けて損失関数への寄与度を調節している。

- (3) それぞれの手法の組み合わせについて5分割交差検証によって分類性能の評価を行い、分類性能の最も高い組み合わせを利用して、収録単語以外の単語のラベルを推定する⁴。分類性能を測る評価基準として、F-尺度 (F-measure)、適合率 (precision)、再現率 (recall) のそれぞれについて、加重平均 (weighted average)、マイクロ平均 (micro average)、マクロ平均 (macro average) の3つの方法で集計した計9個の指標を用いる⁵。
- (4) 推定されたラベルをもとに、再度人手で極性辞書に収録する単語を選定する。具体的には、固有名詞を除外し、経済関連単語のみを抽出する。

表4は、各組み合わせにおける分類性能の評価の結果をまとめたものであり、初期値を変えて各モデルの推定と予測を5回試行した際の平均値が示されている。また、各指標の最も高い数値は太字で表示している。なお、SVMについては、最も分類性能の高かった多項式カーネルの結果のみを記載している。表4から、総じてSVMとSkip-gramの組み合わせの分類性能が最も高いことがわかる。また、学習モデルによって異なるものの、総じてSkip-gramとfastTextはGloVeよりも分類性能が高い傾向にあった。このほか、次元は大きいほど分類性能が高まる傾向にあるが、400次元と800次元とではほぼ同程度、もしくはわずかに800次元の方が良いことがわかる。これらの結果から、収録単語以外の単語のクラスラベル推定における組み合わせには、SVMと800次元のSkip-gramを採用する⁶。そして、改めてアノテーション済みの3,000語を用いて学習し、新聞記事データに含まれる全単語のクラスラベルを推定する。コーパス内に極性値を持たない単語が含まれていることが想定されることから、最終的に辞書には収録されないものを除外するために、ニュートラルのクラスラベルを含めた3クラス分類の分類器を構築している。ポジティブあるいはネガティブのクラスラベルが付与された単語からアノテーション済みの単語を除外したのち、経済や景気と関連のある単語及び固有名詞を再度筆者2人によって選定した。推定された各クラスの単語と追加収録単語の内訳については、表5の示すとおりである。この結果、追加収録する単語数は、ポジティブが39語、ネガティブが152語の計191語となった。

表1のアノテーションの多数決結果からニュートラルを除外した単語と、追加収録単語を合わせて、最終的な景気単語極性辞書の収録単語数は874語となった。その内訳は表6に示すと

³ なお、one-versus-one法による分類精度と比較したところ、総じてone-versus-rest法による分類精度の方が高いことを確認している。

⁴ ハイパーパラメータのチューニングに関しては、付録Aに詳細を記載している。

⁵ 加重平均では、各クラスのサンプルサイズの逆数で重み付けしている。F-尺度の加重平均は、適合率と再現率の加重平均を用いて算出されている。

⁶ 構築時のSkip-gramの次元以外のパラメータについては、ウィンドウサイズは5、最小頻度は5、負例サンプリングは5である。

学習モデル	埋め込み手法	次元	F-measure			Precision			Recall		
			weighted	micro	macro	weighted	micro	macro	weighted	micro	macro
MLR	Skip-gram	100	0.823	0.820	0.688	0.828	0.820	0.676	0.820	0.820	0.705
		200	0.847	0.844	0.733	0.851	0.844	0.717	0.844	0.844	0.752
		400	0.859	0.858	0.751	0.861	0.858	0.746	0.858	0.858	0.759
		800	0.857	0.857	0.746	0.858	0.857	0.745	0.857	0.857	0.748
	GloVe	100	0.798	0.792	0.652	0.808	0.792	0.634	0.792	0.792	0.678
		200	0.809	0.806	0.669	0.815	0.806	0.658	0.806	0.806	0.686
		400	0.816	0.813	0.686	0.823	0.813	0.672	0.813	0.813	0.706
		800	0.825	0.824	0.694	0.827	0.824	0.692	0.824	0.824	0.702
	fastText	100	0.819	0.816	0.681	0.823	0.816	0.671	0.816	0.816	0.693
		200	0.845	0.844	0.730	0.848	0.844	0.721	0.844	0.844	0.741
		400	0.854	0.853	0.749	0.857	0.853	0.744	0.853	0.853	0.758
		800	0.863	0.863	0.759	0.863	0.863	0.762	0.863	0.863	0.757
SVM	Skip-gram	100	0.863	0.865	0.750	0.863	0.865	0.766	0.865	0.865	0.739
		200	0.872	0.873	0.769	0.872	0.873	0.788	0.873	0.873	0.758
		400	0.872	0.876	0.763	0.871	0.876	0.805	0.876	0.876	0.734
		800	0.878	0.881	0.777	0.880	0.881	0.815	0.881	0.881	0.754
	GloVe	100	0.835	0.839	0.701	0.834	0.839	0.729	0.839	0.839	0.683
		200	0.843	0.843	0.723	0.848	0.843	0.735	0.843	0.843	0.723
		400	0.845	0.851	0.718	0.846	0.851	0.764	0.851	0.851	0.691
		800	0.852	0.860	0.727	0.853	0.860	0.792	0.860	0.860	0.686
	fastText	100	0.852	0.853	0.738	0.853	0.853	0.747	0.853	0.853	0.734
		200	0.860	0.861	0.754	0.861	0.861	0.767	0.861	0.861	0.748
		400	0.870	0.875	0.764	0.871	0.875	0.815	0.875	0.875	0.730
		800	0.876	0.878	0.781	0.877	0.878	0.801	0.878	0.878	0.769
FFNN	Skip-gram	100	0.813	0.805	0.684	0.835	0.805	0.668	0.805	0.805	0.729
		200	0.853	0.854	0.737	0.855	0.854	0.745	0.854	0.854	0.736
		400	0.862	0.863	0.749	0.861	0.863	0.763	0.863	0.863	0.738
		800	0.862	0.865	0.750	0.862	0.865	0.772	0.865	0.865	0.734
	GloVe	100	0.768	0.746	0.643	0.824	0.746	0.604	0.746	0.746	0.743
		200	0.827	0.832	0.689	0.825	0.832	0.718	0.832	0.832	0.668
		400	0.825	0.829	0.686	0.823	0.829	0.706	0.829	0.829	0.671
		800	0.822	0.825	0.689	0.824	0.825	0.702	0.825	0.825	0.685
	fastText	100	0.816	0.808	0.689	0.832	0.808	0.668	0.808	0.808	0.726
		200	0.849	0.853	0.733	0.849	0.853	0.763	0.853	0.853	0.711
		400	0.853	0.857	0.733	0.853	0.857	0.768	0.857	0.857	0.711
		800	0.859	0.862	0.747	0.859	0.862	0.772	0.862	0.862	0.731
LightGBM	Skip-gram	100	0.825	0.834	0.667	0.823	0.834	0.711	0.834	0.834	0.641
		200	0.839	0.848	0.697	0.839	0.848	0.753	0.848	0.848	0.664
		400	0.836	0.845	0.693	0.835	0.845	0.756	0.845	0.845	0.656
		800	0.829	0.841	0.677	0.832	0.841	0.760	0.841	0.841	0.634
	GloVe	100	0.809	0.815	0.646	0.807	0.815	0.674	0.815	0.815	0.631
		200	0.806	0.818	0.637	0.804	0.818	0.702	0.818	0.818	0.602
		400	0.815	0.820	0.663	0.816	0.820	0.690	0.820	0.820	0.652
		800	0.806	0.810	0.651	0.807	0.810	0.677	0.810	0.810	0.641
	fastText	100	0.829	0.833	0.686	0.828	0.833	0.717	0.833	0.833	0.667
		200	0.841	0.852	0.700	0.844	0.852	0.785	0.852	0.852	0.654
		400	0.830	0.837	0.690	0.829	0.837	0.733	0.837	0.837	0.663
		800	0.845	0.852	0.716	0.848	0.852	0.762	0.852	0.852	0.692

表 4 極性分類の結果

	ポジティブ	ニュートラル	ネガティブ	計
全単語数	672	637,077	2,386	640,135
単語リストとの重複単語数	290	2,280	430	3,000
重複排除後の単語数	382	634,797	1,956	637,135
追加収録単語数	39	—	152	191

表 5 推定された各クラスの単語と追加収録単語の内訳

	ポジティブ	ネガティブ	計
名詞	198	369	567
動詞	85	113	198
形容詞	38	71	109
計	321	553	874

表 6 景気単語極性辞書の収録単語の内訳

	ポジティブ	ネガティブ
名詞	売上拡大*, 好況, 金融緩和 人手不足, リスクオン*	貸し渋り, 信用収縮, 生産調整 賃金低下*, デフレマインド*
動詞	上向く, 活気づく, 儲かる 持ち直している*, 和らぐ	落ち込む, 買い控える, 下押す 長引く, 冷え込む
形容詞	底堅い, 粘り強い, 力強い 望ましい, いち早い	重い, 不透明さ*, 程遠い 根深い, 険しい

表 7 景気単語極性辞書の収録単語のサンプル. 注: *印は追加収録単語.

おりである⁷. また, 表 7 には収録単語のサンプルとして, 各品詞についてポジティブ単語とネガティブ単語がそれぞれ 5 語記載されている. なお追加収録された単語は (*) で示されている.

6 景気単語極性辞書の応用例: 経済ニュース指数の作成

本章では, 構築した景気単語極性辞書の使用例として, 経済ニュース指数を作成し, その有用性を検討する⁸. まず 1980 年 10 月から 2021 年 12 月までの日経新聞記事の全ての朝刊記事に

⁷ ファイナンスのドメインを対象とする (Ito et al. 2018) の金融専門極性辞書は特定のイベントや人名等の固有名詞や極性が曖昧な単語が含まれることから収録単語の総数は 19,630 語と景気単語極性辞書の収録単語よりも大きい. このうち 547 語が景気単語極性辞書の収録単語と重複しており, 金融専門極性辞書の収録単語総数を分母とする重複単語の比率は 2.8%となっている.

⁸ 経済テキストデータを指数化する方法として, 本稿のように極性辞書を利用する方法のほかに機械学習を利用する方法がある. 例えば, (山本, 松尾 2016), (風戸 他 2019), (Goshima et al. 2021) では内閣府の景気ウォッチャー調査の景気判断理由集を教師データとして機械学習によって指数を直接構築しているが, 本稿では辞書に基づく指数に限定して分析を行う.

おける日ごとに出現する辞書収録単語を

$$\text{Score}_t = \frac{\text{Positive Words}_t - \text{Negative Words}_t}{\text{Positive Words}_t + \text{Negative Words}_t} \times 100 \quad (1)$$

に従って数え上げることで、(センチメント) スコアを算出する。ここで、 Positive Words_t と Negative Words_t はそれぞれ、 t 日に配信された新聞記事中に出現したポジティブな単語の数とネガティブな単語の数を表す。ただし、休刊日のスコアは前日と同じ値を用いることとする。日次単位で計測したスコアは不規則変動が多く、景気の転換点を判定するような場合には平滑化が望ましい。本稿では、初期値を $\text{News Index}_1 = \text{Score}_1$ とし、 $t \geq 2$ について

$$\text{News Index}_t = \text{Score}_t \times \alpha + \text{News Index}_{t-1} \times (1 - \alpha) \quad (2)$$

で計算されるスコアの指数平滑移動平均を経済ニュース指数として定義した。ここで、 $\alpha = 2/(n+1)$ は平滑化係数であり、過去 1 か月間の変動が特に重視されるように、 $n = 30$ と設定した。また、特定のドメインを考慮しない一般的な極性辞書である (Takamura et al. 2005) の単語感情極性値対応表を用いて、同様の経済ニュース指数を作成し、景気連動性に関して、景気単語極性辞書による経済ニュース指数との比較も行う⁹。なお、単語感情極性値対応表には 55,125 語が収録されており、そのうちの 0.72% にあたる 397 語が景気単語極性辞書と重複している。

図 1 は、2 つの辞書を用いて構築した経済ニュース指数の月末値の推移を示している。ただし実線が経済ニュース指数、灰色の領域が内閣府によって認定された景気基準日付に基づく景気後退期の期間である¹⁰。図 1 の (a) からは、景気単語極性辞書から作成された経済ニュース指数がバブル崩壊や IT バブル崩壊、世界金融危機等の景気後退期において低下した様子が観察される。また、2020 年の新型コロナウイルス拡大初期においても急速に指数が低下しており、景気との連動性の高い経済ニュース指数であることが確認できる。一方で、図 1 の (b) の単語感情極性値対応表から作成された経済ニュース指数については、景気後退期に低下する場合があるものの、(a) と比較して景気との連動性が総じて低いことがわかる。

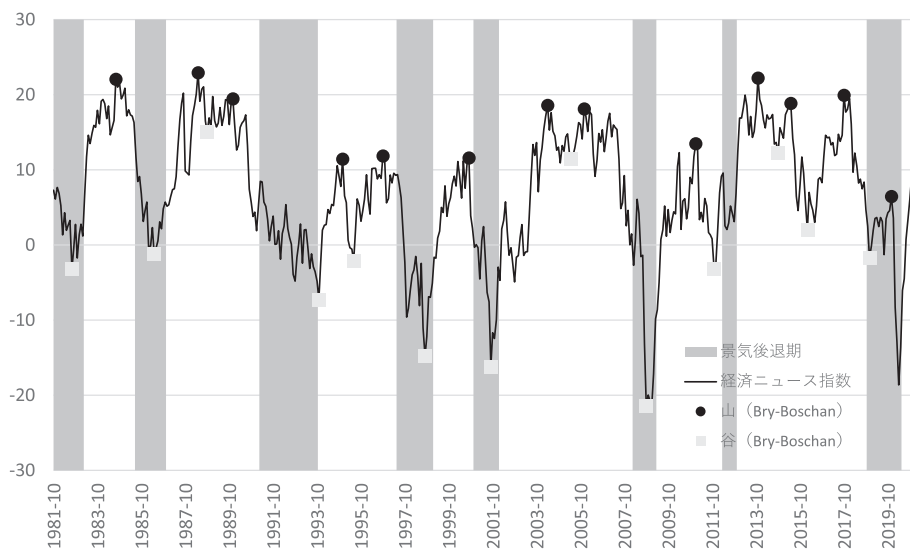
次に、(1) パターン分析、(2) Bry-Boschan 法、(3) 各景気指標との時差相関の 3 つの手法を通じて、経済ニュース指数と景気動向との関係について考察を行う。パターン分析と Bry-Boschan 法は全米経済研究所 (NBER) で開発された景気循環分析の標準的な手法であり、時差相関と合わせて個別系列の景気との連動性評価に用いられることが多い。例えば内閣府が公表する景気

⁹ 単語感情極性値対応表ではラベルが二値ではなく連続値で与えられているため、収録単語を用いた各営業日のスコアを以下の式から算出している。

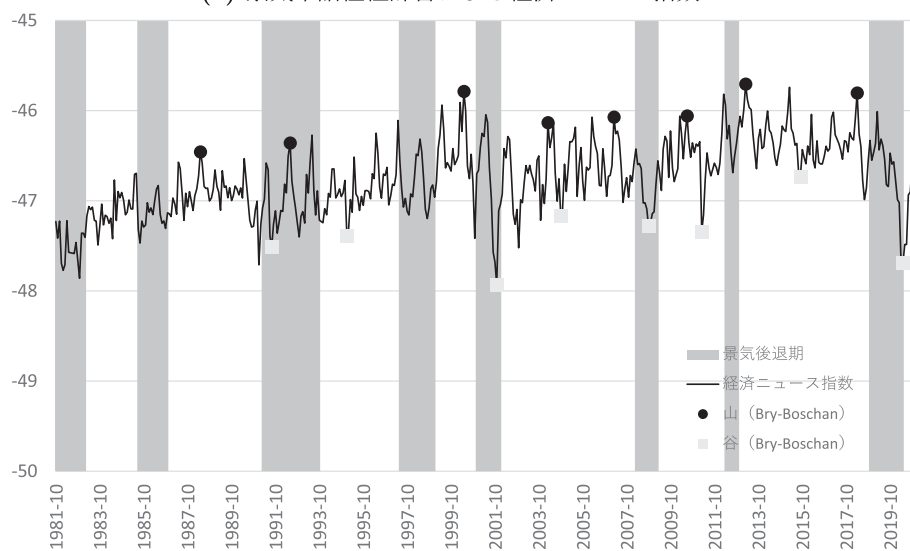
$$\text{Score}_t = \frac{\text{Labels}_t}{\text{Words}_t} \times 100.$$

ここで、 Labels_t と Words_t はそれぞれ、 t 日に配信された新聞記事中に出現した単語の極性値の総和と単語の数を表す。なお、単語を数え上げる際には、重複を許している。

¹⁰ 景気基準日付 (山と谷) については、内閣府のウェブページ (<https://www.esri.cao.go.jp/jp/stat/di/hiduke.html>) から取得可能である。



(a) 景気単語極性辞書による経済ニュース指数

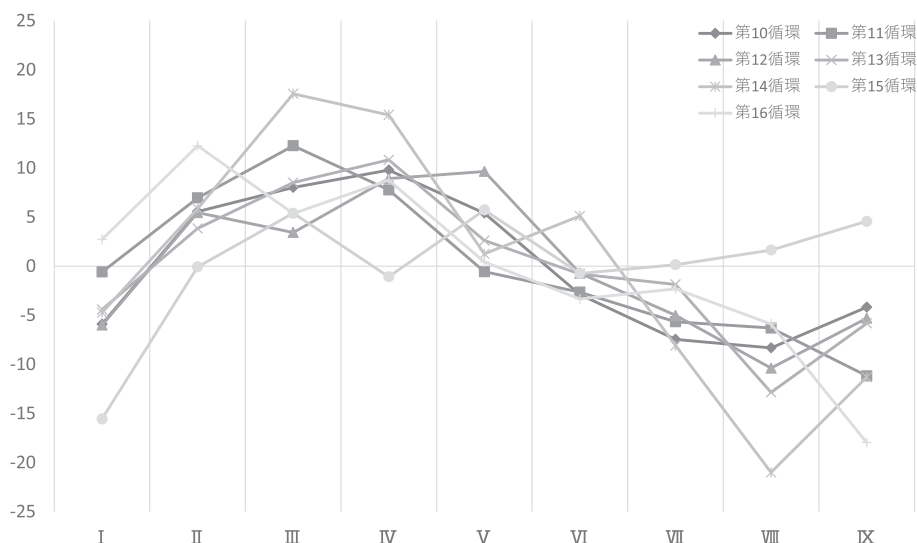


(b) 単語感情極性値対応表による経済ニュース指数

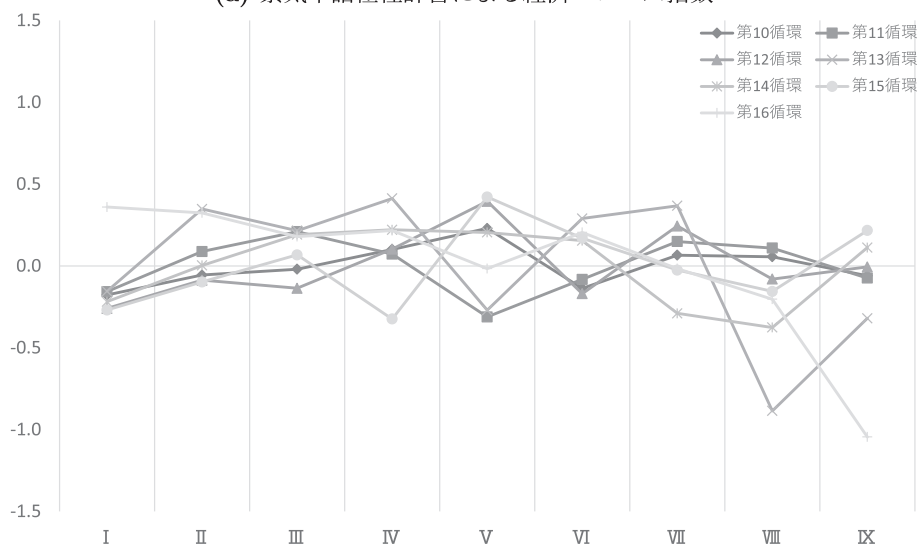
図 1 経済ニュース指数の概観

動向指数の採用系列の変更の際にもこれらの手法が利用されている (内閣府 2015). まずパターン分析は, 対象となる景気の 1 循環を 9 ステージに分割し, 系列の挙動をステージごとに要約する手法であり, 複数の循環の動向を比較することで, 景気基準日付に対する先行遅行関係の

安定性を調べる事ができる¹¹。なお経済ニュース指数が計算できる標本期間と対応させるため、ここでは戦後の景気循環のうち第10循環から第16循環まで（1983年2月から2020年5月まで）の7つの景気循環を分析対象としている。図2は、2つの極性辞書を用いて作成した



(a) 景気単語極性辞書による経済ニュース指数



(b) 単語感情極性値対応表による経済ニュース指数

図2 パターン分析

¹¹ パターン分析の詳細な手順については、付録Bを参照されたい。

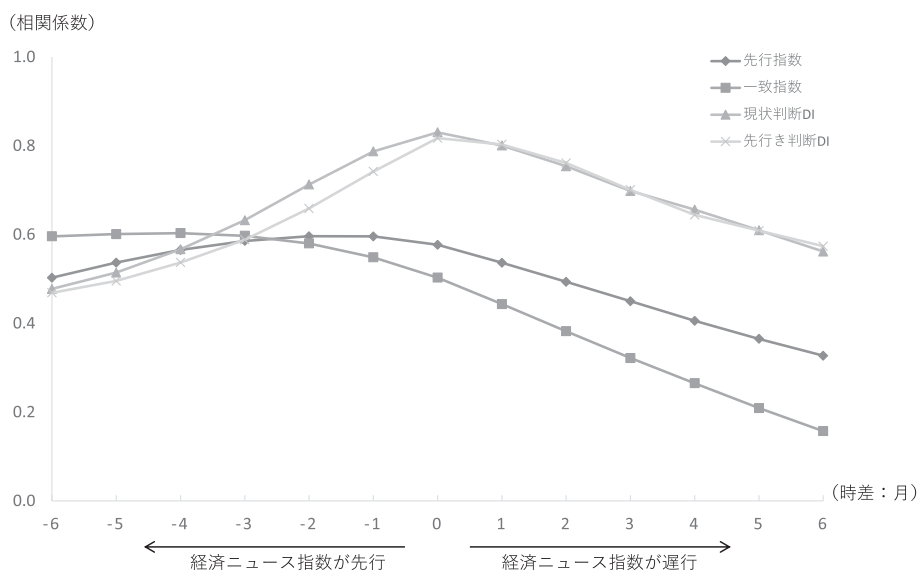
景気指標	公表者	公表頻度	利用系列	期間	内容
景気動向指数	内閣府	月次	CI 先行指数 CI 一致指数	1985/1-2020/12	景気の現状把握及び将来予測の目的で、鉱工業生産指数など重要かつ景気に敏感に反応する複数の系列の変化を集計して作成された合成指標
景気ウォッチャー調査			現状判断 DI 先行き判断 DI	2002/1-2020/12	地域の景気に関連の深い動きを観察できる立場にある人々に対するサーベイ調査から作成された景気動向判断のための指標
全国企業短期経済観測調査 (日銀短観)	日本銀行	四半期	大企業-製造業-実績 大企業-製造業-予測 大企業-非製造業-実績 大企業-非製造業-予測	1983/9-2020/12	日本銀行が金融政策の適切な運営に必要な全国の企業動向把握のために実施する企業のサーベイ調査から作成された指標
生活意識に関する アンケート調査			景況感-前年比 景況感-予想	2006/6-2020/12	生活者の現状に対する生活実感や、金融・経済環境変化が生活者の意識や行動に与える影響を把握する目的のサーベイ調査から作成された指標

表 8 各景気指標のまとめ

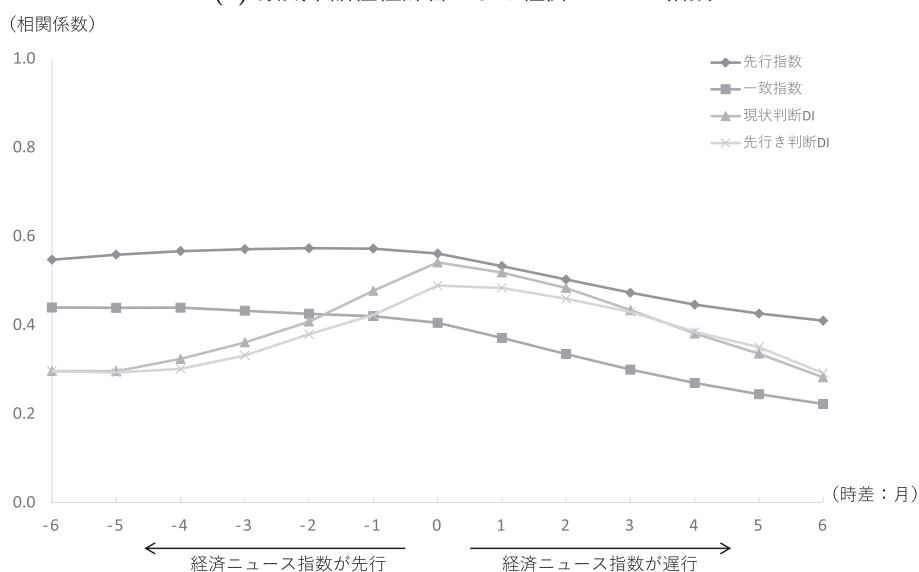
ニュース指数に対して、パターン分析を適用し、7つの景気循環の各ステージ（横軸）における標準化された値（縦軸）を図示したものである。

パターン分析においては、第5ステージが景気の山に対応するが、図2の(a)が示すように、景気単語極性辞書による経済ニュース指数の最大値の多くは第3ステージと第4ステージに位置している。また、第9ステージが景気の谷に対応するが、同じ経済ニュース指数の最小値は概ね第8ステージに位置している。経済ニュース指数の極値が、内閣府による景気の山谷に対して1から2ステージ程度前に観察されることから、景気の拡大局面と後退局面に先行して変動していることがわかる。つまり景気単語極性辞書による経済ニュース指数には、将来の景気動向の予測に有益な情報が含まれていることを示唆している。一方、図2の(b)では、単語感情極性値対応表による経済ニュース指数は、景気単語極性辞書による経済ニュース指数に比べて、景気との関連性は弱く、特に景気拡大局面の先行性を見出すことはできない。

Bry-Boschan 法とは、山と谷との間隔を5か月以上とする、1循環の長さを15か月以上とするなどの制約を課した上で、移動平均等の変換と組み合わせて景気の山谷を判定するアルゴリズムである(Bry and Boschan 1971)。Bry-Boschan 法から計算された系列の山谷のタイミングと景気基準日付との先行遅行関係を調べることで、景気循環における系列の挙動の特徴を把握することができる。図1の(a)に示されている黒丸と四角は景気単語極性辞書による経済ニュース指数にBry-Boschan 法を用いることによって計算された山と谷である。図中の景気後退期は景気基準日付による山から谷までの期間として定義されているため、灰色の領域が始まる点が山、終わる点が谷に対応している。まず景気単語極性辞書による経済ニュース指数の振動回数は景気基準日付の景気循環より多いものの、そのタイミングは概ね景気基準日付と一致している。また、指数の山と谷は、両者ともに政府が認定した景気循環より早めに転換点が来ている



(a) 景気単語極性辞書による経済ニュース指数

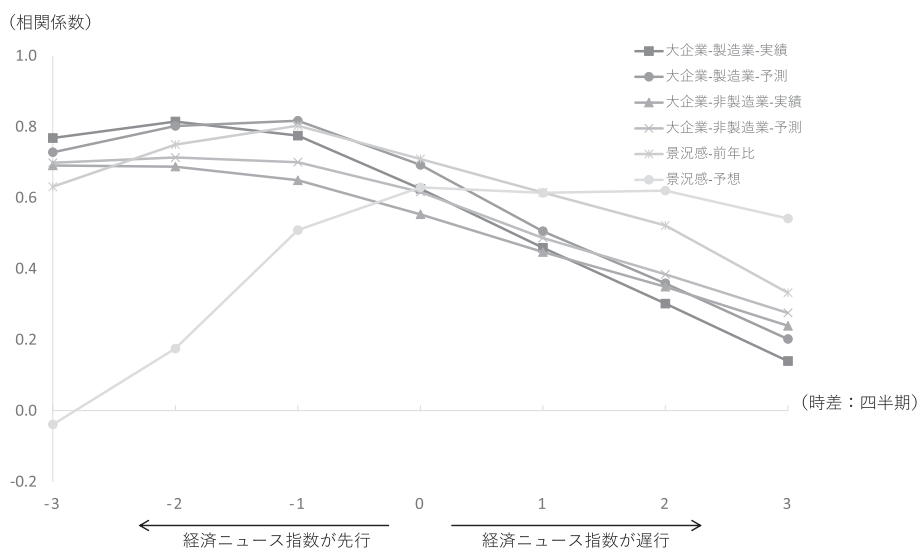


(b) 単語感情極性値対応表による経済ニュース指数

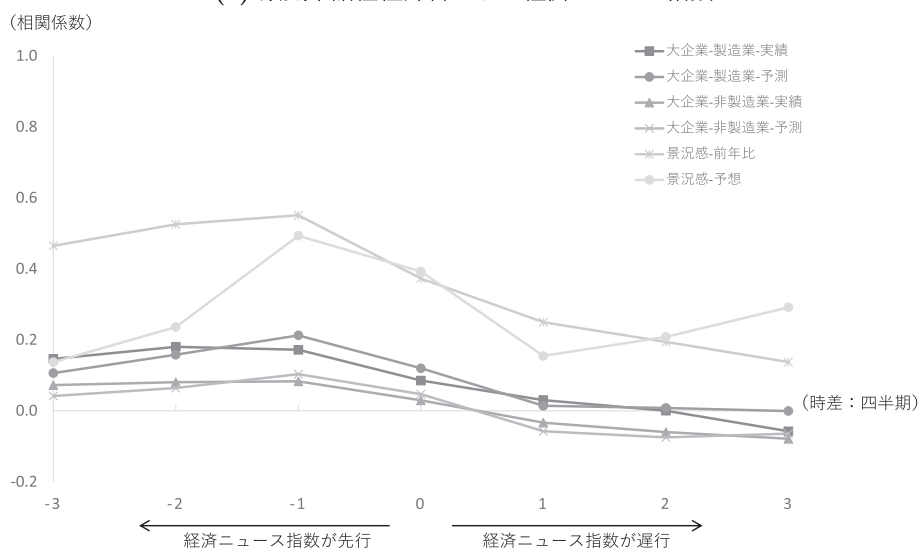
図 3 内閣府が公表する景気指標との時差相関

ことがわかる．一方で，図 1 の (b) の単語感情極性値対応表による経済ニュース指数については特に前半の期間において，山と谷のタイミングが景気基準日付と一致していないことが多い．

最後に，内閣府及び日本銀行によって公表されている各景気指標と経済ニュース指数の先行



(a) 景気単語極性辞書による経済ニュース指数



(b) 単語感情極性値対応表による経済ニュース指数

図 4 日本銀行が公表する景気指標との時差相関

遅行関係を，時差相関係数を用いて考察する．時差相関係数は比較する2つの変数のうち，一方をラグ変数とした場合の相関係数として計算される．複数のラグについてこのような時差相関係数を計算することによって，両系列の先行遅行関係を分析することができる．比較対象となる各景気指標については表8にまとめられている．

図3と図4には，経済ニュース指数と各景気指標との時差相関係数が示されている．ここで

縦軸は相関係数、横軸は時差（月/四半期）であり、横軸の左方向は経済ニュース指数が先行する方向、右方向は経済ニュース指数が遅行する方向を示している。図3の(a)からは、景気単語極性辞書による経済ニュース指数は内閣府が公表している各景気指標と高い相関があり、景気動向を的確に捉えていることが分かる。また景気動向指数に対しては先行性が見られ、先行指数とは-2か月、一致指数とは-4か月で相関係数が最も大きい値を取ることから、景気動向をいち早く捕捉できる可能性を示唆している。景気ウォッチャー調査に対しては、先行性は見られないものの同時点で相関係数が0.8と非常に高い値を取っており、景気との連動性が高いことがわかる。また、図4の(a)からは、景気単語極性辞書による経済ニュース指数は日本銀行が公表している各景気指標とも高い相関があることがわかる。特に日銀短観に対しては、-3四半期または-2四半期で相関係数が最も大きい値を取ることから、明確な先行性が観察される。生活意識に関するアンケート調査に対しては、景況感-前年比とは-1四半期、景況感-予想とは同時点で相関係数が最大となっている。さらに、これらの指標は集計から発表までのラグがあるため、経済ニュース指数はナウキャストの目的でも有用であるといえる。一方で、図3の(b)と図4の(b)からは、単語感情極性値対応表による経済ニュース指数は総じて各景気指標との相関が低いことがわかる。以上の結果から、本稿で構築した景気単語極性辞書のマクロ経済分析における有用性を確認することができた。特に景気単語極性辞書を用いることで、景気動向をいち早く捕捉できる経済ニュース指数が作成できる可能性が示唆された。

7 まとめ

本稿では、経済研究における日本語テキスト分析に寄与することを目的として、マクロ経済分野に適した極性辞書の構築と景気動向分析への応用例を示した。具体的には、日本経済新聞の新聞記事コーパスから選択された単語について、複数人の専門家によるアノテーションを行い、景気単語極性辞書を構築した。また、景気単語極性辞書を用いて作成された経済ニュース指数は、景気動向を迅速に捕捉する上で有用であることが確認された。景気単語極性辞書の少量コーパスへの適用や景気動向分析以外への応用は今後の課題である。

謝 辞

本稿の作成に当たっては、リサーチアシスタントの上田佳祐氏には多岐にわたり助力を頂いた。ここに記して感謝したい。

参考文献

- Baker, S. R., Bloom, N., and Davis, S. J. (2016). “Measuring Economic Policy Uncertainty.” *The Quarterly Journal of Economics*, **131** (4), pp. 1593–1636.
- Bojanowski, P., Grave, E., Joulin, A., and Mikolov, T. (2017). “Enriching Word Vectors with Subword Information.” *Transactions of the Association for Computational Linguistics*, **5**, pp. 135–146.
- Bry, G. and Boschan, C. (1971). *Cyclical Analysis of Time Series: Selected Procedures and Computer Programs*. National Bureau of Economic Research, Inc.
- 五島圭一, 高橋大志 (2017). 株式価格情報を用いた金融極性辞書の作成. 自然言語処理, **24** (4), pp. 547–577. [K. Goshima and H. Takahashi (2017). Building a Sentiment Dictionary for News Analytics using Stock Prices. *Journal of Natural Language Processing*, **24** (4), pp. 547–577.].
- Goshima, K., Ishijima, H., Shintani, M., and Yamamoto, H. (2021). “Forecasting Japanese Inflation With a News-based Leading Indicator of Economic Activities.” *Studies in Nonlinear Dynamics & Econometrics*, **25** (4), pp. 111–133.
- Hansen, S. and McMahon, M. (2016). “Shocking Language: Understanding the Macroeconomic Effects of Central Bank Communication.” *Journal of International Economics*, **99** (S1), pp. S114–S133.
- Ito, T., Sakaji, H., Tsubouchi, K., Izumi, K., and Yamashita, T. (2018). “Text-Visualizing Neural Network Model: Understanding Online Financial Textual Data.” In Phung, D., Tseng, V., Webb, G., Ho, B., Ganji, M., and Rashidi, L. (Eds.), *Advances in Knowledge Discovery and Data Mining. PAKDD 2018. Lecture Notes in Computer Science, Springer, vol. 10939*, pp. 247–259.
- Kaji, N. and Kitsuregawa, M. (2007). “Building Lexicon for Sentiment Analysis from Massive Collection of HTML Documents.” In *Proceedings of the 2007 Joint Conference on Empirical Methods in Natural Language Processing and Computational Natural Language Learning*, pp. 1075–1083.
- Kalamara, E., Turrell, A., Redl, C., Kapetanios, G., and Kapadia, S. (2022). “Making Text Count: Economic Forecasting Using Newspaper Text.” *Journal of Applied Econometrics*, **37** (5), pp. 896–919.
- Kanayama, H. and Nasukawa, T. (2006). “Fully Automatic Lexicon Expansion for Domain-oriented Sentiment Analysis.” In *Proceedings of the 2006 Conference on Empirical Methods in Natural Language Processing*, pp. 355–363.

- 風戸正行, 黒崎哲夫, 五島圭一 (2019). 日本銀行による景気判断のトーン分析. 金融研究所デイスカッションペーパーシリーズ 2019-J-16, 日本銀行. [M. Kazato et al. (2019). Nihonginko niyoru Keikihandan no Ton Bunseki. The IMES Discussion Paper Series 2019-J-16, Bank of Japan.].
- 小林のぞみ, 乾健太郎, 松本裕治, 立石健二, 福島俊一 (2005). 意見抽出のための評価表現の収集. 自然言語処理, **12** (3), pp. 203–222. [N. Kobayashi et al. (2005). Collecting Evaluative Expressions for Opinion Extraction. Journal of Natural Language Processing, 12 (3), pp. 203–222.].
- Liu, B. (2012). *Sentiment Analysis and Opinion Mining*, Vol. 5. Morgan and Claypool Publishers.
- Loughran, T. and McDonald, B. (2011). “When Is a Liability Not a Liability? Textual Analysis, Dictionaries, and 10-Ks.” *Journal of Finance*, **66** (1), pp. 35–65.
- Mikolov, T., Sutskever, I., Chen, K., Corrado, G. S., and Dean, J. (2013). “Distributed Representations of Words and Phrases and their Compositionality.” In *Advances in Neural Information Processing Systems*, pp. 3111–3119.
- 内閣府 (2015). 第 16 回景気動向指数研究会 議事要旨・資料. 内閣府経済社会総合研究所景気統計部. [Cabinet Office, Government of Japan (2015). Dai 16 Kai Keiki Doko Shisu Kenkyukai Giji Yoshi Shiryo. Naikakufu Keizai Shakai Sogo Kenkyusho Keiki Tokeibu.].
- 日本銀行 (2020). 経済・物価情勢の展望 (2020 年 10 月 30 日) . [Bank of Japan (2020). Outlook for Economic Activity and Prices (October 30, 2020).].
- 王悠介, 須合智広, 高橋耕史, 松村浩 (2021). 位置情報データによる経済活動のナウキャストイング. ワーキングペーパーシリーズ No. 21-J-2, 日本銀行. [Y. Oh et al. (2021). Ichijohodeta niyoru Keizaikatsudo no Naukyasutingu. Working Paper Series No. 21-J-2, Bank of Japan.].
- Pennington, J., Socher, R., and Manning, C. (2014). “GloVe: Global Vectors for Word Representation.” In *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing*, pp. 1532–1543.
- Qiu, G., Liu, B., Bu, J., and Chen, C. (2009). “Expanding Domain Sentiment Lexicon through Double Propagation.” In *Proceedings of the 21st International Joint Conference on Artificial Intelligence*, pp. 1199–1204.
- 佐藤貴俊, 高村大也, 奥村学 (2016). 分散表現を用いた単語の感情極性抽出. 情報処理学会, 2016-NL-228 12 号, pp. 1–6. [T. Sato et al. (2016). Bunsanhyogen wo Mochiita Tango no Kanjo Kyokusei Chushutsu. IPSJ SIG Technical Report, 2016-NL-228 (12), pp. 1–6.].
- 佐藤敏紀, 橋本泰一, 奥村学 (2017). 単語分かち書き辞書 mecab-ipadic-NEologd の実装と情報検索における効果的な使用方法の検討. 言語処理学会第 23 回年次大会発表論文集, pp. 875–878. [T. Sato et al. (2017). Tango Wakachigaki Jisho mecab-ipadic-NEologd no Jisso

- to Jouho Kensaku niokeru Kokateki na Shiyo Hoho no Kento. The 23th Annual Meeting of the Association for Natural Language Processing, pp. 875–878.].
- Takamura, H., Inui, T., and Okumura, M. (2005). “Extracting Semantic Orientations of Words using Spin Model.” In *Proceedings of the 43rd Annual Meeting of the Association for Computational Linguistics*, pp. 133–140.
- Tang, D., Wei, F., Qin, B., Zhou, M., and Liu, T. (2014a). “Building Large-Scale Twitter-Specific Sentiment Lexicon : A Representation Learning Approach.” In *Proceedings of the 25th International Conference on Computational Linguistics: Technical Papers*, pp. 172–182.
- Tang, D., Wei, F., Yang, N., Zhou, M., Liu, T., and Qin, B. (2014b). “Learning Sentiment-Specific Word Embedding for Twitter Sentiment Classification.” In *Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pp. 1555–1565.
- Turney, P. D. (2002). “Thumbs Up or Thumbs Down? Semantic Orientation Applied to Un-supervised Classification of Reviews.” In *Proceedings of the 40th Annual Meeting of the Association for Computational Linguistics*, pp. 417–424.
- 山本裕樹, 松尾豊 (2016). 景気ウォッチャー調査の深層学習を用いた金融レポートの指数化. 第 30 回人工知能学会全国大会論文集. 3L3OS16a2, [Y. Yamamoto and Y. Matsuo (2016). Sentiment Summarization of Financial Reports by LSTM RNN model with the Japan Economic Watcher Survey Data. The 30th Annual Conference of the Japanese Society for Artificial Intelligence, 3L3OS16a2.].

付録

A ハイパーパラメータの選択

本稿では, FFNN 以外の各モデルについて, 訓練データに対する 5 分割交差検証とグリッドサーチによって, ハイパーパラメータを決定した後, 全ての訓練データを利用してモデルを学習している. 具体的には, 訓練データを 5 分割し, そのうちの 1 つの副標本を検証データとし, 残りのデータで学習を行う. その際, 検証データを用いた分類評価指標 (F-尺度の加重平均) をパラメータのグリッド上の各点で計算する. この作業を別の副標本を検証データとした場合にも繰り返し, 5 回の分類評価指標の平均値が最も高いパラメータの値を選択する. チューニングしたハイパーパラメータとその探索範囲は表 9 のとおりである.

FFNN については隠れ層を 3 層とし, 各層の次元数は入力ベクトルのサイズの 3 倍, 4 倍, 1

学習モデル	ハイパーパラメータ	範囲
MLR	C	{0.001, 0.01, 0.1, 1, 10, 100}
	ノルム	{L1, L2}
SVM (線形カーネル)	C	{0.001, 0.01, 0.1, 1, 10, 100}
SVM (ガウシアンカーネル)	C	{0.001, 0.01, 0.1, 1, 10, 100, 1000}
	γ	{0.001, 0.01, 0.1, 1}
SVM (多項式カーネル)	C	{0.0001, 0.001, 0.01, 0.1, 1, 10, 100, 1000}
	次数	{2, 3, 4}
	γ	{0.001, 0.01, 0.1, 1}
SVM (シグモイドカーネル)	C	{0.0001, 0.001, 0.01, 0.1, 1, 10, 100, 1000}
	γ	{0.001, 0.01, 0.1, 1}
LightGBM	学習器の数	{10, 20, 40, 80}
	木の最大の深さ	{4, 8, 16, 32}
	葉の最大数	{4, 8, 16, 32}
	学習率	{1, 0.3, 0.1, 0.03, 0.01, 0.003, 0.001}

表 9 各モデルのハイパーパラメータとその探索範囲

倍としている。加えて、各層にドロップアウトを採用している。最適化手法は Adam を採用し、訓練データのうち 20% を検証データとしたエポックの早期打ち切りを行っている。

B パターン分析

本稿で用いたパターン分析の手順を以下に示す。

- (1) まず、景気の 1 循環は内閣府で定めた景気基準日付の谷から次の谷までとする。1 循環の最初の谷、山、2 番目の谷のそれぞれについて、該当する月とその前後 1 か月を合わせた計 3 か月を、順に第 1 ステージ、第 5 ステージ、第 9 ステージと定める。
- (2) 景気の拡張期である第 1 ステージと第 5 ステージの間を 3 等分し、第 2、第 3 及び第 4 ステージに分割する。
- (3) 景気の後退期である第 5 ステージと第 9 ステージの間を 3 等分し、第 6、第 7 及び第 8 ステージに分割する。
- (4) 9 個のステージに該当する期間ごとに平均値を計算する。等分する際に割り切れなかった月については、 $2/3$ または $1/3$ の重みを付することで調整する。
- (5) 各ステージの値を 1 循環分 (9 ステージ) の平均値で割って 100 を乗じることで標準化する。
- (6) 以上の計算を 1 循環ごとに繰り返し、ステージを横軸にした図に重ねて表示する。

略歴

五島 圭一：2012 年慶應義塾大学経済学部卒業。2017 年東京工業大学大学院総合理工学研究科知能システム科学専攻博士後期課程修了。博士（工学）。日本銀行金融研究所，早稲田大学商学部講師，東京大学エコノミックコンサルティング株式会社などを経て 2022 年 5 月より株式会社三菱 UFJ 銀行にて産業リサーチ業務に従事。早稲田大学産業経営研究所招聘研究員を兼任。

新谷 元嗣：1991 年大阪大学経済学部卒業。2000 年米イエール大学で博士号（経済学）取得。慶應義塾大学商学部講師，米ヴァンダービルト大学経済学部教授，東京大学先端科学技術研究センター教授を経て，2018 年より東京大学大学院経済学研究科教授，2020 年より東京大学エコノミックコンサルティング株式会社アドバイザー（兼任）。

高村 大也：1997 年東京大学工学部計数工学科卒業。2000 年同大学院工学系研究科計数工学専攻修了（1999 年はオーストリアウィーン工科大学にて研究）。2003 年奈良先端科学技術大学院大学情報科学研究科博士課程修了。博士（工学）。2003 年から 2010 年まで東京工業大学精密工学研究所助教。2006 年にはイリノイ大学にて客員研究員。2010 年から 2016 年まで同准教授。2017 年から 2021 年 3 月まで同教授。2017 年より産業技術総合研究所人工知能センター知識情報研究チーム研究チーム長。計算言語学，自然言語処理を専門とし，特に機械学習の応用に興味を持つ。

（2022 年 3 月 1 日 受付）

（2022 年 6 月 1 日 再受付）

（2022 年 7 月 17 日 再々受付）

（2022 年 7 月 20 日 採録）