

## 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

Sentiment Summarization of Financial Reports by LSTM RNN model with the Japan Economic Watcher Survey Data

山本 裕樹 \*1\*2 松尾 豊 \*2  
Yuhki Yamamoto Yutaka Matsuo

\*1野村証券株式会社 金融工学研究センター

Nomura Securities co., ltd. Financial Engineering &amp; Technology Research Center

\*2東京大学工学系研究科技術経営戦略学専攻

Graduate School of Engineering, the University of Tokyo

For many investors, it is very time-consuming to read financial and economic reports issued by financial institutions, central banks and governments. This paper introduces the method that calculates an economic sentiment index from texts in a financial report by using the machine learning technique. The index can assist investors in interpreting the overall opinion presented in a financial news article or report. The model is investigated with Recurrent Neural Network (RNN) language model by learning the relations between sentences and economic sentiments in the Japanese Economic Watcher Survey data. An economic sentiment index is obtained by applying this sentiment estimate model to all sentences in a financial report. To evaluate effectiveness, we calculated the indices of the monthly economic reports issued by the Bank of Japan and the Government of Japan and confirmed that these indices have positive and significant correlation with the Japanese stock index (Nikkei 225).

## 1. はじめに

政府や日銀のマクロ経済レポートから証券会社のアナリストレポートまで、膨大な数の金融レポートが日々発行されている。投資家はこれらのレポートから情報を得て、自らの投資判断を下すために多くの時間を費やす必要がある。

数値データであれば、データ数が膨大でも合計や平均等で集計することによって簡単に有用な情報を得ることが出来る。実際、GDP や日経平均指数のような経済指標はこのような数値データの集計値と言え、金融経済の実態を知るために広く活用されている。一方、テキストデータの場合には数値データのような簡単な集計は難しく、GDP や日経平均のような広く用いられている指数は筆者の知るところ存在しない。投資家は時間的な制約から、膨大な金融レポートの極一部しか読むことが出来ず、大量のテキストデータが十分に活用されていないのが現状である。これらのテキストには金融経済の情勢を把握し将来を予測するのに有用な情報が多くあると推察される。

本稿では、深層学習を用いた機械学習モデルによって自動的にテキストの景気センチメントを判別することで、金融レポートのテキストを低コストかつ高速に数値(センチメント指数と呼ぶ)化して集計する手法を提案する。機械学習モデルとしては、近年発展の著しい深層学習の一手法である、Long Short Term Memory(LSTM) ユニットを持つ Recurrent Neural Network(RNN) モデルを用いた。学習データには内閣府が調査・公表している景気ウォッチャー調査を用いた。景気ウォッチャー調査は景気に関するテキストとセンチメントが紐付いた約 20 万サンプルの大規模データであり、世界的にも希少なデータと言える。景気ウォッチャー調査のテキストからその景気センチメント(ポジティブ/ネガティブ)を予測するタスクを RNN で学習したところ、約 95% の高い精度で文書の景気センチメントを判別させることが可能となった。景気ウォッチャー調査データと深層学習を用いて、景気に関する文書の高精度な

センチメント推定に成功した点は本稿の成果の一つと考えられる。得られたセンチメント推定モデルを使って政府や日銀が発行する月次レポートのセンチメントを推定し指数化したところ、得られた指数の変動はマクロの景気変動の動きと良く連動していることを確認できた。実際、日経平均との相関を算出したところ、既に投資指標として広く用いられている日銀短観や景気ウォッチャー指数と比べても高い値を示した。元が純粋なテキストデータであることを考えると、この結果は大きな進展と考えられ、今後の金融テキストデータの集計、活用の活発化が期待される。本手法によってテキスト情報の利用が促進されることで、金融市場がより効率化され、それによって資金を必要としている人や企業、政府に適切に投資が集まる事が期待できる。また、特に機関投資家に比べて金融レポートを読む時間や知識の制約が大きい個人投資家にとって有効なツールになり得ると考えられ、投資家間の情報格差を埋める役割も期待される。

なお、本稿の手法で得られた指数は、野村 AI 景況感指数として野村証券から公表されている [山本 15]。指数の解釈や投資情報としての活用方法についての詳細はこちらも参照されたい。

## 2. 関連研究

金融・経済分野におけるテキストデータを用いた関連研究について述べる。

ニュースやレポート等のテキストデータを投資に活用する試みは多くある。特に近年では本稿と同様に文章のセンチメントを解析し、株価等のマーケットデータとの連動性を示すことを報告するものが多数ある。例えば、[Heston 14]、[石島 13] ではニュース記事のセンチメントを単語の出現頻度と辞書から評価し、得られた指標が株価と相関を持つことを示した。また、[迫村 13] では Twitter のテキストに対して同様のセンチメント分析を行っている。[諏訪部 15] では、四季報や決算議事録やアナリストレポートのセンチメントを予め用意した辞書で評価し、それを元に株式を売買することで統計的に有意な正のり

連絡先: 山本 裕樹, 野村証券 (株) 金融工学研究センター,  
yuhki.yamamoto@nomura.com

ターンを得られるとしている。これらはいずれもテキスト解析では単語を数え上げる手法 (BOW) を用いているが、3.2.2 で述べる通り、文脈情報が無視されていることで精度が落ちている可能性がある。本稿で用いた RNN の様な深層学習を用いたセンチメント分析としては、映画レビューテキストのセンチメントを推定した研究は多く、特に LSTM を用いることで高い精度が得られることが報告されている [Hong 15, Tai 15]。また RNN を用いた金融関連テキストの分析としては、ニュース記事の要約を試みた [Chen 15] がある。一方、RNN を用いた金融テキストのセンチメント分析を行った研究は筆者の知る限りでは無い。景気ウォッチャー調査を学習データに用いることで、金融テキストのセンチメントを高精度に推定した点は本稿の新規性の一つと考えている。

景気ウォッチャー調査を用いた研究としては [岡崎 15] が挙げられ、語彙の出現頻度や共起ネットワークと各月の DI との関係性を調べ、DI が変化した理由を語彙レベルに分解し要因を明らかにする試みが成されている。テキストに限らず、ビッグデータから景気指数を算出したり投資への活用を試みる研究として、検索データを利用した [Edelman 12, Xiong 15]、POS データを使用した [渡辺 13] 等が挙げられる。

### 3. 景気センチメント判別機の学習

テキストデータの指数化に用いたモデルの学習について説明する。

#### 3.1 学習データ：景気ウォッチャー調査

学習データには内閣府が調査・公表している景気ウォッチャー調査 [wat] を用いた。景気ウォッチャー調査とは内閣府が行う景気動向に関するアンケート調査で、タクシー運転手や小売店の店主等、景気に敏感な人達 (景気ウォッチャー) に景気に対する判断 (5 段階、現状・先行き別) とその理由 (テキストデータ、現状・先行き別) をアンケートしている。結果は毎月公表され、景気判断の集計値が PDF 等で公表されるのに加えて、全件のテキストデータと景気判断が CSV で利用可能となっている。本稿では、この景気判断理由のテキストから、景気判断を予測する教師有り学習によって、景気に関する文章からそのポジティブ/ネガティブ度合を自動で判断する機械を学習した。景気ウォッチャー調査データの回答別件数を表 1 にまとめた。解答は景気の実況判断と先行き判断に分かれており、いずれも中央の「変わらない」を中心に凡そ対象に分布している。実際の回答サンプルや質問項目の詳細については [wat] を参照された。

表 1: 景気ウォッチャー調査の回答別件数

	現状判断	先行き判断
良い	1,757	1,787
やや良い	20,406	22,976
変わらない	41,810	47,482
やや悪い	20,053	19,612
悪い	6,373	5,971
合計	90,399	97,828

#### 3.2 景気センチメント判別機の構築

用いた機械学習モデルの概要と推定精度について述べる。

#### 3.2.1 LSTM RNN モデル

モデルは LSTM RNN を用いた。RNN は時系列データのような連続した系列をニューラルネットワークで解析する手法であり、長さが異なる系列への適用も可能であるため、本稿のような自然言語の解析に向いている。RNN では、前の層の中間層を次の層の入力ベクトルの一部として用いるため、過去の系列を”記憶”することができる。言語の解析ではこれによって文脈の解析が可能となる。しかし、RNN の活性化関数に以下のような単純な tanh 関数を用いた場合、誤差の勾配が消失または発散してしまう問題があり、長期の系列の学習が難しいことが知られている。

$$h_t = \tanh(Wx_t + Uh_{t-1} + b)$$

ここで、 $x_t$  は  $t$  時点の入力、 $h_{t-1}$  は  $t - 1$  時点の隠れ層、 $W$ 、 $U$ 、 $b$  はそれぞれの重み行列、切片である。LSTM [Hochreiter 97, Gers 00] はこの RNN の勾配消失・発散問題を解決するために用いられ、近年、言語のセンチメント判別タスクにおいても高い性能を発揮することが報告されている [Hong 15, Tai 15]。LSTM では上式の  $h_t$  の代わりに、下式で表される入力ゲート ( $i_t$ )、忘却ゲート ( $f_t$ )、出力ゲート ( $o_t$ ) を用いることで、それぞれ入力、前の隠れ層の重み、出力の重みを調節して記憶することが出来、これによって長期の記憶を可能としている。

$$i_t = \sigma(W^{(i)}x_t + U^{(i)}h_{t-1} + b^{(i)})$$

$$f_t = \sigma(W^{(f)}x_t + U^{(f)}h_{t-1} + b^{(f)})$$

$$o_t = \sigma(W^{(o)}x_t + U^{(o)}h_{t-1} + b^{(o)})$$

$$u_t = \tanh(W^{(u)}x_t + U^{(u)}h_{t-1} + b^{(u)})$$

$$c_t = i_t \circ u_t + f_t \circ c_{t-1}$$

$$h_t = o_t \circ \tanh(c_t)$$

ここで  $\sigma$  はシグモイド関数、 $\circ$  は要素間の積を表す。なお、出力層のセンチメント予測は [Tai 15] を参考に最終時点の隠れ層のみを用いて行った。

#### 3.2.2 モデル精度の検証

表 2 にセンチメント判別モデルの精度について纏めた。表中の”二値”は「良くなる」、「やや良くなる」を 1、「悪くなる」、「やや悪くなる」を 0 とした二値判別モデル、”回帰”は「悪くなる」= -2 ~ 「良くなる」= +2 を回帰したモデルである。表中の数値は”二値”では正答率、”回帰”では平均二乗誤差 (mse) を示している。”現状”は景気ウォッチャーデータの現状判断のみ、”先行き”は先行き判断のみ、”現状・先行き”は両者を用いた。”RNN”は LSTM ユニットを用いた RNN モデル、”BRNN”はさらに Bidirectional を用いた。”tf-idf”は tf-idf 法による素性を用いて、ロジスティック回帰 (二値) または線形回帰 (回帰) でモデル構築を行った。いずれの結果も全データの 90% を学習データ、10% をテストデータとして、テストデータの精度を表記している。また、パラメーターの選択は学習データの中からさらに 10% を検証データに用いて行った。RNN の主なパラメーターは中間層の次元=250、世代=3、バッチサイズ=50、ドロップアウト=0.3 である。

景気ウォッチャーデータの中の”現状”の景気判断のみを用いた二値判別では LSTM RNN モデル (RNN) が、それ以外では Bidirectional LSTM RNN (BRNN) が最も精度が高くなった。tf-idf はいずれのモデルでも RNN に比べて大きく精度が

劣る。これは文脈情報が失われるため、例えば「悪くはならない」、「良くなるとは思えない」といった係り受けが複雑な文章で正解できないためと考えられる。この結果から、景気ウォッチャーデータを用いた景気センチメントの判別では、RNNのような文脈情報を考慮したモデル化を用いることが重要であると考えられる。

表 2: センチメント推定精度の比較

	RNN	BRNN	tf-idf
二値 (現状)	<b>0.930</b>	0.928	0.859
二値 (先行き)	0.939	<b>0.946</b>	0.869
二値 (現状・先行き)	0.940	<b>0.947</b>	0.878
回帰 (現状・先行き)	0.346	<b>0.340</b>	0.536

## 4. 金融・経済レポートの指数化

### 4.1 指数化の方法

ここでは月例経済報告を例に、指数化の方法を説明する。月例経済報告は日本政府が毎月発行する経済レポートである。景気動向に関する日本政府の公式見解を示しており、閣議資料等で用いられる重要な文書である。図表はほぼ無く、1998年1月から2015年12月までの平均で毎月17,036文字、219文から成るレポートである。

本稿では読点(“.”)までを一つの文として、各文ごとにモデルでセンチメントを評価し、それを各月ごとに平均してその月の指数値とした。モデルは平均の意味が解釈し易い点、および「変わらない」を含む全てのサンプルで学習できる点を考慮して、回帰モデルを用いた。回帰モデルでは、景気が「悪い」=-2~「良い」=+2として学習しているため、指数の最小値は-2、最大値は+2である。

### 4.2 得られた指数

得られた指数の時系列(1998年1月から2015年12月)を図1に示した。まず政府の月例経済報告から得られた指数と、日銀の金融経済月報から得られた指数が非常によく似た動きを示していることは望ましい結果と言える。いずれの文書も各時点の景気動向に書かれたレポートであるので、本来、似たような景気センチメントを持つべきであるが、もしノイズが大きい手法で指数化を行った場合、異なる文書がこのような高い連動性を偶然持つとは考え難く、本手法の有効性を示す結果と言える。また、両指数ともにマクロ経済の動きを良く反映しており、「ITバブルの崩壊」、「リーマンショック」、「東日本大震災」、「アベノミクス」といった大きなイベントによって株価と連動して上下している。

表3は他の時系列との相関を示している。表の右上半分が相関係数、左下半分がそのt値である。まず、政府の月例経済報告から得られた指数( $Idx^{GOV}$ )と日銀の金融経済月報から得られた指数( $Idx^{BOJ}$ )は、共に投資指標として良く知られた日銀短観(TAN D.I.)、四半期毎)や景気ウォッチャー指数(WAT D.I.)と非常に強い相関を示した。日銀短観と景気ウォッチャー指数がともに大規模なアンケート調査から得られていることを考えると、非常に低コストで算出可能な本指数の意義はそれだけで大きいと考えられる。さらに、投資指標としての活用可能性を探る上では重要となる日経平均株価(N225)との相関係数は、本稿の手法で算出した指数( $Idx^{BOJ}$ )が最も高く、日銀短観や景気ウォッチャー指数を上回る株価との連動性を示した。この結果は本稿のようにテキストデータを指数化することで、

実際に投資情報として活用できる可能性があることを示唆していると言える。

表 3: 指数間の相関係数 (t 値 \ 係数)

	$Idx^{GOV}$	$Idx^{BOJ}$	TAN	WAT	N225
$Idx^{GOV}$	-	0.901	0.821	0.843	0.387
$Idx^{BOJ}$	30.4	-	0.822	0.782	<b>0.476</b>
TAN D.I.	21.1	21.1	-	0.740	0.294
WAT D.I.	22.9	18.4	15.1	-	0.464
N225	6.1	<b>7.9</b>	4.5	7.7	-

## 5. まとめ

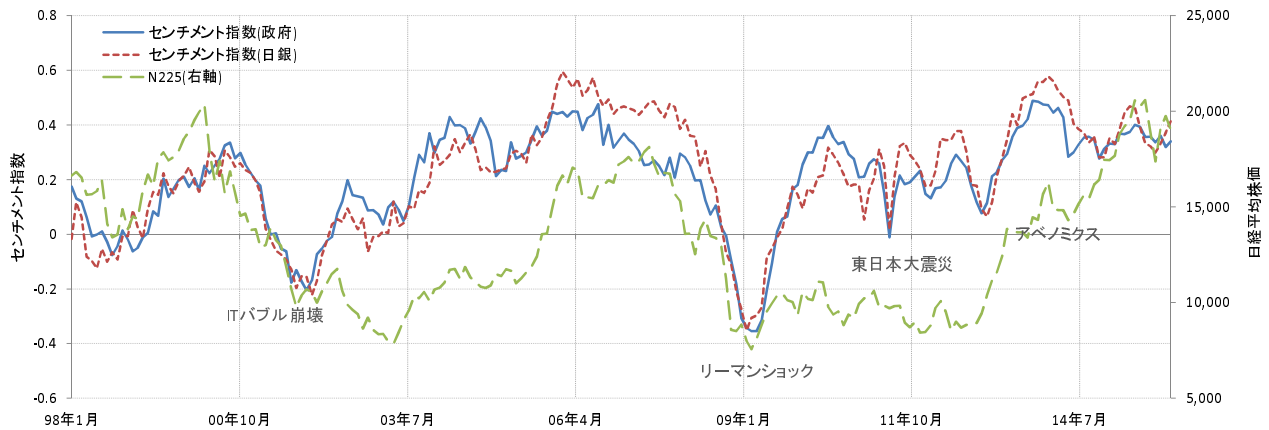
景気ウォッチャー調査を学習データに用いて、景気に関する文章のセンチメントを自動判別できる機械学習モデルを開発した。LSTM RNNモデルによって95%近い精度で景気ウォッチャー調査の文章のセンチメントを推定することが可能である。得られたモデルを用いることで、景気ウォッチャー調査以外の景気に関する文章に対しても、センチメントの判別を行うことができる。一つの応用例として、日本政府が発行する「月例経済報告」と日本銀行が発行する「金融経済月報」の景気センチメントを指数化した。得られた指数は景気の高、谷を反映した動きを示しており、日経平均株価との相関は日銀版で0.47(t値=7.9)であった。この日経平均との相関係数は、投資指標として一般によく用いられる日銀短観や景気ウォッチャー指数と比較しても高く、投資情報として活用できる可能性がある。また、日銀短観や景気ウォッチャー指数が大規模なアンケート調査によって得られていることを考えると、コストの観点からも有用である。

本稿の手法は様々なテキストデータに応用可能であり、これまで投資家の時間、知識、言語等の制約から十分な活用が難しかった、膨大なテキストデータを投資情報として活用できる可能性がある。今後の展望として、まず得られた指数による株価や経済指標、政府・日銀の政策の予測可能性を調べることで、投資情報としての活用することが考えられる。これには、即時性を高めることが重要であるが、金融機関が発行する日次のレポート、ニュース記事やSNS等から指数(ナウキャスト)を算出することが考えられる。これによってより株価の予測性が高く、投資情報として有効な指数を開発できると期待される。

## 参考文献

- [Chen 15] Chen, K. Y., Liu, S. H., Chen, B., Wang, H. M., Jan, E. E., Hsu, W. L., and Chen, H. H.: Extractive Broadcast News Summarization Leveraging Recurrent Neural Network Language Modeling Techniques, *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, Vol. 23, pp. 1322–1334 (2015)
- [Edelman 12] Edelman, B.: Using Internet Data for Economic Research, *Journal of Economic Perspectives*, Vol. 26, pp. 189–206 (2012)

図 1: 得られたセンチメント指数の時系列



[Gers 00] Gers, F. A., Schmidhuber, J., and Cummins, F.: Learning to forget: Continual prediction with LSTM, *Neural computation*, Vol. 12, No. 10, pp. 2451–2471 (2000)

[Heston 14] Heston, , L., S., and Sinha, N. R.: News versus Sentiment: Comparing Textual Processing Approaches for Predicting Stock Returns, *Robert H. Smith School Research Paper* (2014)

[Hochreiter 97] Hochreiter, S. and Schmidhuber, J.: Long short-term memory, *Neural computation*, Vol. 9, No. 8, pp. 1735–1780 (1997)

[Hong 15] Hong, J. and Fang, M.: Sentiment Analysis with Deeply Learned Distributed Representations of Variable Length Texts, *stanford.edu* (2015)

[Tai 15] Tai, K. S., Socher, R., and Manning, C. D.: Improved semantic representations from tree-structured long short-term memory networks, *arXiv preprint arXiv:1503.00075* (2015)

[wat] 景気ウォッチャー調査 - 内閣府, [http://www5.cao.go.jp/keizai3/watcher/watcher\\_menu.html](http://www5.cao.go.jp/keizai3/watcher/watcher_menu.html)

[Xiong 15] Xiong, R., Nichols, E., and Shen, Y.: Deep Learning Stock Volatility with Google Domestic Trends, *arXiv:1512.04916v3* (2015)

[岡崎 15] 岡崎 陽介, 敦賀 智裕: ビッグデータを用いた経済・物価分析について, 日本銀行レポート・調査論文 (2015)

[山本 15] 山本 裕樹, 水門 善之, 木下 智夫: 人工知能で政府・日銀の景況感を指数化する, NOMURA マクロ・エコノミクス・インサイト (2015)

[諏訪部 15] 諏訪部 貴嗣: データ革命と株式運用戦略, 証券アナリストジャーナル (2015)

[石島 13] 石島 博, 数見 拓朗, 前田 章: 日次データを用いた市場センチメント・インデックスの構築と株価説明力の分析, 第 11 回人工知能学会金融情報研究会資料 (2013)

[渡辺 13] 渡辺 広太, 渡辺 努: スキャナーデータを用いた日次物価指数の計測, *CARF Working Paper Series* (2013)

[迫村 13] 迫村 光秋, 和泉 潔: Twitter テキストマイニングによる経済動向分析, 第 9 回人工知能学会 ファイナンスにおける人工知能応用研究会 (2013)