

2021 年 卒業論文

論文題名：深層学習を用いた少量データセットによる  
景気動向分析

指導教員：藤井章博

法政大学理工学部

応用情報工学科

学籍番号： 18X3145

氏名：森田 彩星

確認印

# 目次

1. 序論.....	3
1-1. 研究背景 .....	3
1-2. 研究目的 .....	4
2. 関連研究.....	5
3. 分析対象.....	6
4. 分析手法.....	8
5. 要素技術.....	11
5-1. LSTM .....	11
5-1-1. BiLSTM .....	14
5-1-2. ATTENTION .....	15
5-2. BERT .....	16
5-3. GPT-2 .....	17
5-4. 形態素解析.....	18
5-5. 分散表現 .....	18
6. 結果.....	19
6-1. 深層学習手法の精度 .....	19
6-2. GPT-2 による文章生成.....	23
6-3. 繰り返し学習 .....	24
7. 考察.....	26
7-1. 深層学習手法の精度に対する考察.....	26

7-2.	GPT-2 による文章生成に対する考察 .....	27
7-3.	繰り返し学習に対する考察 .....	29
8.	結論 .....	30
	参考文献 .....	31
	謝辞 .....	34

# 1. 序論

## 1-1. 研究背景

現在、緊急事態宣言の発出と解除、新規感染数の増加と減少に合わせて経済活動の制限と緩和が繰り返されていて[\[1\]](#)、現在の日本の景気は、新型コロナウイルス感染症の影響から引き続き厳しい状況にある[\[2\]](#).

ここで、景気とは一般的に経済に関する消費財やサービスの売買動向などの経済活動の状況[\[3\]](#)を指し、景気動向を把握することは、企業業績や企業売上に関係する。それに対して、現在の感染症の影響などの景気動向の乱れがある際に、迅速に正確な景気動向を確かめる必要があり、より分析精度が高くなるモデルの研究が必要となる。

現在は、サポートベクターマシンやロジスティック回帰、ランダムフォレストなどの機械学習手法を用いた分析が普及しつつあり、その中でも、人間の神経細胞の仕組みを模倣したニューラルネットワークを用いた深層学習手法では、画像処理、音声認識、自然言語処理などの多岐にわたる分野で大きな成果を生み出している[\[4\]](#)。自然言語処理の分野においては、2018年に BERT という汎用モデルが開発され、今では自然言語処理の分野における標準的なモデルとしての地位を確立した。

ただ、少量のデータセットからでも高い性能を示すことができる BERT を用いた分析をする際においても、ある程度の量のデータセットが分析に必要な

## 1-2. 研究目的

本研究では、景気文章を対象とした深層学習手法による景気動向予測を行う。しかし、日本語の文章を対象とした分析をする際に、大量のデータセットの取得が容易ではない場合がある。この場合、深層学習を用いた分析において、データ数が少ないと精度の向上に期待ができない。

ここで、データを新たに生成することができれば深層学習を用いた分析をする際の精度が上がると仮定して、本研究では、データ数の水増しをして景気文章の分析をする。ただ問題点として、文章の水増し手法はいまだに確立されていないということである。

この問題に対処するため、**GPT-2** という文章自動生成をして分析対象の文章数を増やすことで、文章自動生成による水増し手法の有効性の検討を行う。

**GPT-2** とは、深層学習を使用した文章自動生成モデルで、**RNN** やマルコフ連鎖を使用した文章自動生成では異なり、あらかじめ大量の語彙で言語パターンを学習したモデルを使用することで、入力した文章以外の単語を含めた、それらしい文章を生成するという特徴がある。下記の「5-3. **GPT-2**」に **GPT-2** の概要を示す。

## 2. 関連研究

景気文章を対象に深層学習を使用して分析した研究として[5]では内閣府が調査・公表している景気ウォッチャー調査の文章を対象に, LSTM と RNN の深層学習モデルを使用して景気センチメントを自動判別する手法を提案し, [6]では日本銀行の景気文章を対象に, RNN モデルによる分析に加えて LDA によるトピック分類をすることで, どのトピックの要素が景気状況に影響を与えているかについての提案を行った.

また GPT-2 を用いた既存研究について, [7]では人の動作データを対象とした将来の動作予測などの非言語データを扱う研究などもあるが, 言語データを扱う研究として, GPT-2 を使用した景気変動について分析した[8]では景気ウォッチャー調査を対象として, 学習用文章から GPT-2 を使用して更に文章数を増やし, 精度の向上の可能性を見出した.

また[9]等では, RNN モデルを使用して景気ウォッチャー調査を初期モデルとして分析を行い, 得られた結果の一部を学習データとしてより良いモデルを学習する方法により, 文章データの経済分析への利用可能性を示した.

### 3. 分析対象

本研究では、景気分析でよく使用されている日本銀行が掲載している全国企業短期経済観測調査(以下、「短観」という.)という統計調査を用いた。

短観とは、全国の企業動向を的確に把握し、金融政策の適切な運営に資することを目的とした統計調査で、全国の約1万社の企業を対象に、四半期毎に実施している[\[10\]](#)。

短観には業況判断 D.I.という企業の経営者が景気の現状・先行きをどのように見ているかを表す指標があり、このD.I.は様々な企業が回答した、景気状況の「良い」「さほど良くない」「悪い」のアンケートから、「良い」の割合を「悪い」の割合で引いたものが数値で表されている。業況判断 D.I.の例を図1に示す。

	大 企 業					
	2021年9月調査		2021年12月調査			
	最近	先行き	最近	変化幅	先行き	変化幅
製造業	18	14	18	0	13	-5
繊維	-8	-8	0	8	2	2
木材・木製品	24	0	19	-5	0	-19
紙・パルプ	19	11	15	-4	0	-15

図1 業況判断 D.I.

また、様々な企業が図1のようなD.I.の数値を予測し、四半期毎に企業のサイトに掲載している。大和総研が掲載している予測文の例[\[11\]](#)を図2に示す。

### 【業況判断 DI (最近)】製造業で供給制約の影響が顕在化

2021 年 10 月 1 日に公表予定の 9 月日銀短観において、大企業製造業の業況判断 DI (最近) は+15%pt (前回調査からの変化幅: +1%pt)、大企業非製造業では+1%pt (同: ±0%pt) を予想する(図表 1)。業況判断 DI の水準は製造業と非製造業の間に大きな差があり引き続き業種間格差が見られるものの、コロナショック以降に顕著であった製造業における業況判断 DI の急速な改善は一服する公算が大きい。

大企業製造業の業況判断 DI は小幅ながらも上昇が続く見込みだが、その内容はこれまでの傾向と異なる。これまで素材業種と加工業種の業況判断 DI は概ね同程度のペースで改善が続けてきたが、9 月短観では商品市況の上昇を主因とした素材業種の業況判断 DI の改善が製造業全体をけん引する見込みだ。具体的には、「石油・石炭製品」や「鉄鋼」といった業種の業況判断 DI が改善を示すとみている。対照的に加工業種に関しては、原材料価格の上昇が収益を圧迫することに加え、供給制約の影響が顕在化しており、業況判断 DI は悪化に転じるとみている。特に、「自動車」に関しては半導体不足や部品調達の難航を受けた生産抑制の影響が顕在化しており、業況判断 DI の悪化を見込む。

大企業非製造業に関しては、緊急事態宣言及びまん延防止等重点措置の対象地域拡大や期間延長を受け「対個人サービス」、「宿泊・飲食サービス」といった業種の業況判断 DI は引き続き低水準で推移するとみている。「小売」は新型コロナウイルス感染症対策の悪影響を受けやすいものの、東京オリンピック・パラリンピック関連需要の拡大が収益を下支えしたとみられる。他方、「卸売」に関しては資源価格の上昇が恩恵として働いたことで業況判断 DI の改善が見込まれる。

図 2 企業の予測文



## 4. 分析手法

本研究では、様々な企業が掲載している業況判断 D.I.の予測文書を対象に BERT と LSTM を用いて 5 クラス分類を行い、予測した分類モデルから景気動向予測を行った。

使用した教師データは、企業が掲載している D.I.の値を予測している文書を使用し、本研究では手動で取得することができた 2008 年 6 月から 2021 年 6 月までの合計 277 個の予測文書を句点(“。”) で区切った文を一文とした合計 760 文で実装をした。また、正解データは日銀が四半期毎に掲載している D.I.の値を使用し、D.I.の値によって 5 段階にラベル付けをした。ラベル付けの範囲に関しては、初めに 3 段階(「良い」「普通」「悪い」)に分け、その中で「良い」と「悪い」を 2 分割した。このようにして 5 段階にすることで、3 クラス分類よりも柔軟な予測をすることができる。以下の表 1 にラベル付けの範囲を示す。

表 1 ラベル付け

ラベル	範囲	文章数
とても良い	$15 \leq X$	299 文
良い	$5 \leq X < 14$	220 文
普通	$-5 \leq X < 5$	109 文
悪い	$-15 \leq X < -5$	66 文
とても悪い	$X < -15$	66 文

ここで使用した全期間の D.I.は-58 以上 25 以下である。

また、本研究では少量データセットを対象とし、深層学習手法による予測精度を向上させるために、GPT-2 による文章生成を行い文章数の水増しをした。

以下の図 3 に、少量データセットで景気動向をより正確に予測するためのプロセスを示す。

- ① 景気文章を GPT-2 に入力し、文章数を水増しする。
- ② 水増しされた文章を、形態素解析ツールを用いて形態素毎に分割する。
- ③ 分割した形態素毎に高次元のベクトル表現を獲得する。
- ④ 入力文章のベクトル表現と正解ラベルを深層学習モデルに入力し、学習済みモデルを作成する。ここで使用したモデルを初期モデルとする。
- ⑤ 元の景気文章を④で作成した初期モデルに入力し、予測ラベルを出力させる。ここで、正解ラベルと予測ラベルが大きく異なる場合、元の文章を削除する。  
  
ここでいう正解ラベルと予測ラベルが大きく異なるとは、初期にラベル付けを行った 3 段階(「良い」「普通」「悪い」)のラベルを使用し、正解ラベルと予測ラベルがこの 3 段階において異なる場合を示す。
- ⑥ 文章数が⑤で減らされた新たな景気文章を対象に再び形態素解析、分散表現を獲得し、深層学習モデルに入力し、学習済みモデルを作成する。  
  
ここで使用したモデルを通常モデルとする。
- ⑦ ⑤, ⑥を繰り返す。

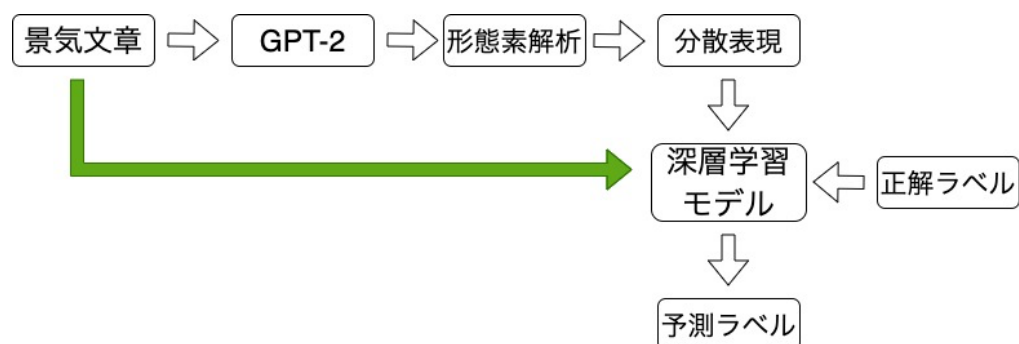


図 3 本研究のプロセス

また、景気動向予測には、文章全文を上記のプロセスで作成した、最後の学習済み分類モデルに入力し、出力された値から期間毎に平均値を取り、その平均値をその期間の予測値とした。

この時にモデルから出力される値は、5 クラス分類モデルを使用したため 5 段階の値に出力される。その後、全期間の予測値を 0~1 の範囲に正規化をし、期間毎に平均値を算出した。

## 5. 要素技術

上記の「4. 分析手法」で使用した要素技術を以下に示す.

### 5-1. LSTM

LSTM(Long Short Term Memory)とは, 時系列データを分析するのに適している深層学習手法で, この LSTM の内部にはループ構造があり, 過去の情報を記憶しながら最新のデータへと更新することができる. LSTM の構造を以下の図 4 に示す.

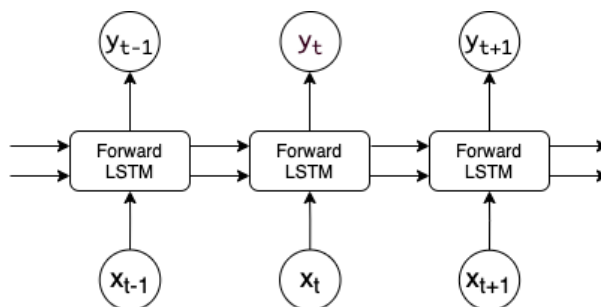


図 4 LSTM の構造

LSTM は, 各時刻で短期記憶を担う  $h_t$  と長期記憶を担う記憶セル  $c_t$  を持ち, 各時刻でそれぞれ入出力を行う. また, LSTM の内部には, 情報の流れを制御する 3 つのゲート(入力ゲート, 出力ゲート, 忘却ゲート)を持ち, 各ゲートは入出力される値に対する重みを算出し, この重みは対応する入出力される値へ掛け合わされる. この重みが LSTM への入力に応じて変化することで, 柔軟な時系列情報の考慮が可能となり, 文脈情報をより長期的に保持することができる. 記憶セル  $c_t$ , 入力ゲート  $i$ , 出力ゲート  $o$ , 忘却ゲート  $f$  の式と, LSTM の内部構造を以下の図 5 に示す.

$$o = \sigma(x_t W_x^{(o)} + h_{t-1} W_h^{(o)} + b^{(o)}) \quad (5.1)$$

$$f = \sigma(x_t W_x^{(f)} + h_{t-1} W_h^{(f)} + b^{(f)}) \quad (5.2)$$

$$g = \tanh(x_t W_x^{(g)} + h_{t-1} W_h^{(g)} + b^{(g)}) \quad (5.3)$$

$$i = \sigma(x_t W_x^{(i)} + h_{t-1} W_h^{(i)} + b^{(i)}) \quad (5.4)$$

$$c_t = f \odot c_{t-1} + g \odot i \quad (5.5)$$

$$h_t = o \odot \tanh(c_t) \quad (5.6)$$

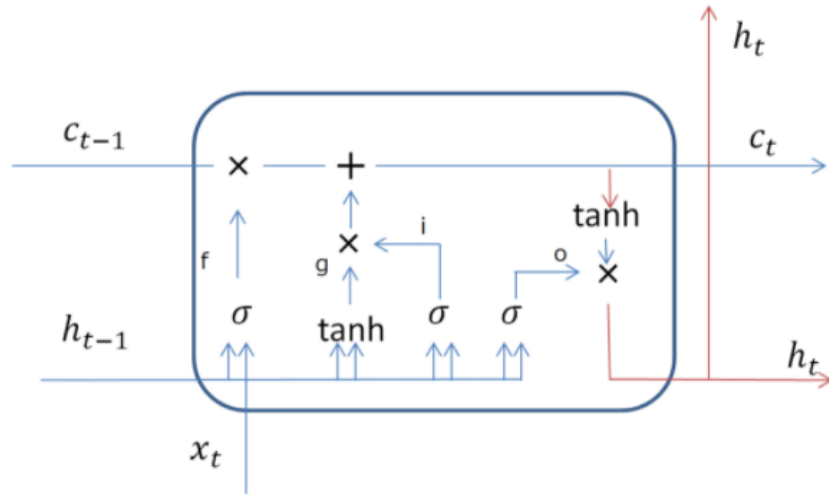


図5 LSTM の内部構造

上の式に使用されている $\odot$ はアダマール積といい、要素ごとの積を計算する。また、 $\sigma$ は sigmoid 関数を表している。

式(5.1)は出力ゲートと呼ばれ、LSTM の出力部分である式(5.6)における  $\tanh(c_t)$  の各要素に対して、それらが次時刻にどれだけ重要かを調整する。

式(5.2)は忘却ゲートと呼ばれ、前時刻の記憶セル  $c_{t-1}$  から、不要な記憶を削除する。また、式(5.3)により、記憶すべき情報を  $\tanh$  により記憶セルに追加する。

式(5.4)は入力ゲートと呼ばれ、式(5.3)の  $g$  に対して新たにゲートを追加する。これにより、 $g$  の各要素が追加する情報として価値があるかを判断する。

以上の 3 つのゲートと、前時刻の記憶セル  $c_{t-1}$  により、式(5.5)の次時刻の記憶セルを計算する[\[12\]](#)。

この LSTM で、文章においても単語の並びを時系列として扱うことで、文章の文脈情報を保持しながら分析を行うことができる。また、LSTM は通常文章の前から後の順方向へ学習を行うが、文章を後から前への逆方向も加えた双方向で学習を行う BiLSTM(双方向 LSTM)や、この BiLSTM に、文章中の重要な単語に着目することができる Attention(注意機構)という手法を加えた BiLSTM-Attention という手法など多岐にわたる可能性がある。

### 5-1-1. BiLSTM

BiLSTM(Bidirectional LSTM)とは、文章の前から後の順方向と、後から前への逆方向の双方向で学習を行う LSTM であり、単方向 LSTM の場合は過去の情報のみを保持するが、未来の情報も加味することで、より文章の文脈情報を理解しやすくなる。

ただ、本研究のような分析対象の文章が完結している場合以外だと、逆方向の処理をすることが困難になる。

BiLSTM の例を以下の図 6 に示す。

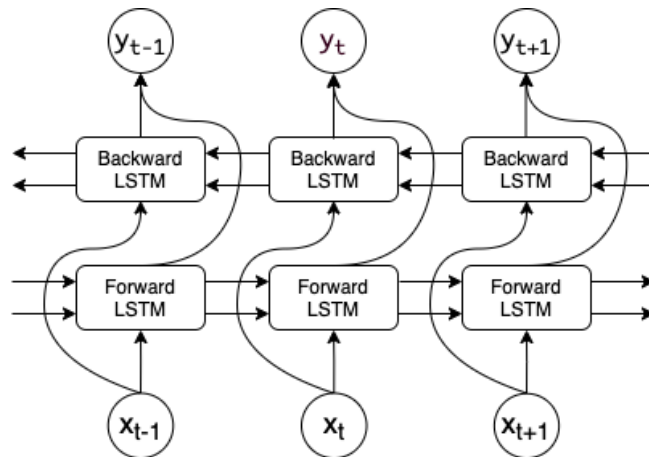


図 6 BiLSTM

### 5-1-2. Attention

Attention とは、連続するデータを扱う際に過去の重要な点に着目する手法で、時系列データから時系列データへ変換を行う seq2seq モデルの改良版となる。seq2seq モデルは、入力文を固定次元の意味ベクトルにして、その固定ベクトルを出力文に変換するが、この固定次元の意味ベクトルは、入力文が長文となると精度が悪くなるという問題点がある。この問題点を改善するために Attention が提案され、seq2seq の固定次元の意味ベクトルに加え、入力単語への参照を行うことで seq2seq よりも長文を処理する際の精度が良くなることが確認された [\[13\]](#)。

この Attention の手法を LSTM に加えることで、LSTM で処理する際の文章中の重要な単語に着目でき、より文脈情報を理解しやすくなる。



## 5-2. BERT

BERT (Bidirectional Encoder Representations from Transformers)とは、2018 年に Google が発表した自然言語処理モデルで、「GLUE(General Language Understanding Evaluation)」という英語圏における様々な自然言語タスクのベンチマークにおいて、発表当初は既存のモデルを上回る性能が確認された。

BERT の構造は、LSTM や RNN(Recurrent Neural Network)の構造とは異なり、Attention のみを利用した Transformer という機械翻訳モデルで構成されている。Transformer には入力した文章を意味ベクトルにする Encoder と、意味ベクトルを文章にする Decoder があり、BERT には Transformer の Encoder が何層も重なって構成されている。これにより、入力した文章中のどの単語に注目すべきかを判断し、文章の文脈を理解することができる。

また BERT の学習には、事前学習とファインチューニングという 2 つの過程があり、事前学習は大量のテキストデータにより学習をさせて言語の基礎理解を学習し、ファインチューニングは少量のテキストデータにより様々なタスクに対応することができる。従って、事前学習済みモデルさえあれば比較的容易に実装をすることができ、分析する際にデータ量が少量しかない場合でも使用できる手法となる[\[14\]](#)。本研究では、日本語版ウィキペディア全文で学習された、東北大学乾研究室の事前学習済み BERT モデル[\[15\]](#)を使用した。

### 5-3. GPT-2

GPT-2 は 2019 年に非営利団体 OpenAI が開発した文章自動生成モデルで、与えられた文章中の単語を元に、逐次的に次の単語を予測することができる。例を図 7 に示す。

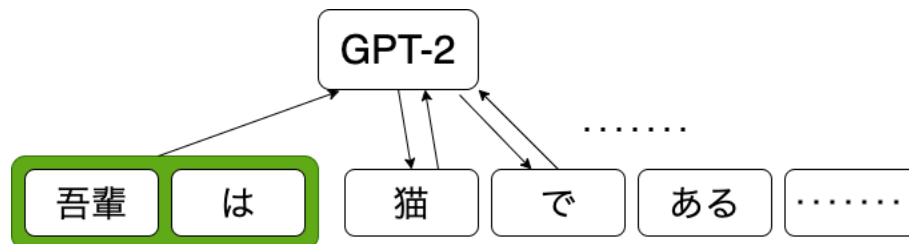


図 7 GPT-2

GPT-2 の構造は, Transformer の Encoder 部を重ねた構造の BERT とは違い, Transformer の Decoder 部を重ねた構造である. Decoder 部は, Masked Multi-Head Attention という, 与えられた文章中の単語だけを使用して確率的に次の単語を予測するアルゴリズムで構成され, この Decoder 部を何層も重ねることにより, 自動文章生成などをすることが可能となる

#### 5-4. 形態素解析

形態素解析とは、文章を形態素という単語の最小単位に分割する手法である。例を図 8 に示す。

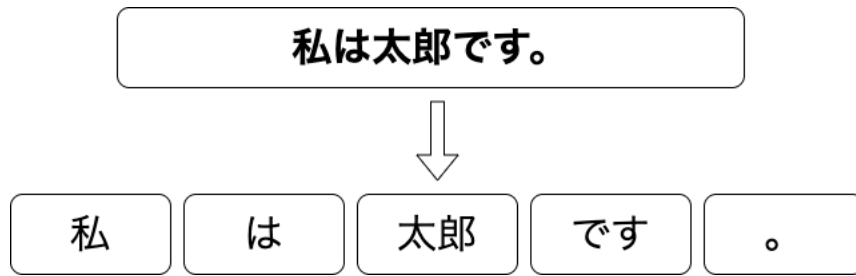


図 8 形態素解析

形態素解析ツールには, MeCab, Juman, Janome などがあるが, 本研究では, LSTM と BERT 共に MeCab を使用した。

#### 5-5. 分散表現

分散表現とは、単語を高次元のベクトルで表現することで, LSTM などのモデルに文章を入力する際, 文章を形態素解析した後に, 分割した単語毎に分散表現を獲得する。例を図 9 に示す。

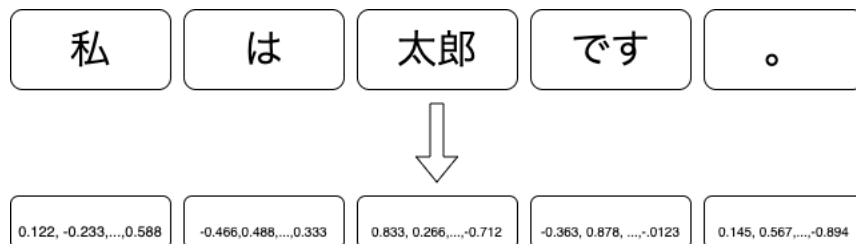


図 9 分散表現

## 6. 結果

上記の「4. 分析手法」と「5. 要素技術」では、それぞれ、本研究の研究内容  
と、本研究に用いた要素技術について述べた。

以下に、上記の研究内容に対する結果を示す。

### 6-1. 深層学習手法の精度

初めに文章を分類する際の深層学習手法による 5 クラス分類の分類精度比較  
を行った。以下の表 2 に結果を示す。

表 2. 深層学習手法の分類精度

分類モデル	分類精度
RNN	0.484
BERT	0.500
TextCNN	0.547
LSTM	0.568
BiLSTM	0.578
BiLSTM-Attention	<b>0.625</b>

表 2 から少量のデータセットでは BiLSTM-Attention が 5 クラス分類をす  
るのに適していることがわかる。ここで、表 2 で使用した分類モデルの  
BiLSTM-Attention による景気動向の予測結果と、比較のための BERT の景気  
動向の予測結果を以下の図 10, 図 11 に示す。

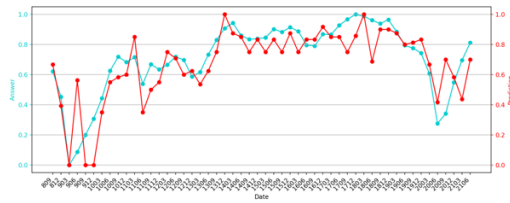


図 10 BiLSTM-Attention による景気動向予測

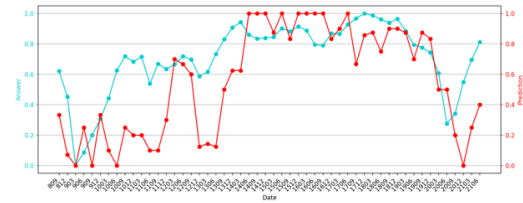


図 11 BERT による景気動向予測

上の図は業況判断 D.I.の実測値(青線)と景気動向を予測した値(赤線)のグラフである。図から BiLSTM-Attention は BERT よりも実測値と予測値の差が大きいように見えるが、BiLSTM-Attention でも実測値と予測値の差が大きい箇所が複数確認できた。

ここで、予測結果の良し悪しを判断する指標として、4つの回帰分析の評価指標と、ある期間から次の期間の値の増減の正解率を用いて、本研究では予測精度とは別に比較を行った。以下に4つの評価指標と増減の正解率の詳細を示す。また、以下に示す式では実測値を  $y$ , 予測値を  $\hat{y}$ , 平均値を  $\bar{y}$  とする。

### ① RMSE

2 乗平均平方根誤差といい、値が小さい方が良いモデルと言える。また、外れ値の影響を受けやすいという特徴がある。

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2}$$

## ② MAE

平均絶対誤差といい, RMSE 同様に値が小さい方が良いモデルと言える.  
ただ, RMSE とは異なり外れ値の影響を受けにくいという特徴がある.

$$\text{MAE} = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i|$$

## ③ AIC

赤池情報量基準といい, モデルの当てはまり度を表す統計量で, 値が小さければ小さいほど良いモデルと言える.

$$\text{AIC} = -2(\log L - k)$$

L: 統計モデルの最大尤度    k: モデルに含まれるパラメーター数

## ④ 決定係数 $R^2$

モデルの当てはまり度合いを表していて, 値は 1 に近いほど良いモデルと言える. ただ, モデルの当てはまりが良くないと値が負になる.

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2}$$

## ⑤ 増減の正解率

本研究では, ある期間から次の期間への変化量が全体の-5% ~ 5%なら「変化なし」とし, それ以外は「上昇」, 「減少」に分類した. 本研究では, この 3 つのラベルを用いて増減の正解率を算出した.

以下の表 3 と表 4 に BiLSTM-Attention と BERT についての評価指標および増減正解率を示す.

表 3 評価指標

分類モデル	RMSE	MAE	AIC	R <sup>2</sup>
BERT	0.291	0.229	-38.128	-0.584
BiLSTM-Attention	<b>0.139</b>	<b>0.103</b>	<b>-61.265</b>	<b>0.637</b>

表 4 増減の正解率

分類モデル	増減の正解率
BERT	0.635
BiLSTM-Attention	<b>0.692</b>

表から, BiLSTM-Attention は全ての指標で BERT を上回っているので, BERT よりも景気動向を予測できていると言える.

## 6-2. GPT-2 による文章生成

次に, GPT-2 による文章生成を行い景気文章の水増しをさせた場合の比較を行った. ここで, 景気文章全文を GPT-2 に入力して, GPT-2 の事前学習済みモデルをファインチューニングさせた. これにより, 入力した文章の特徴を持つ文章を生成することができる.

本研究では, 景気文章全文を入力し, 生成された文章が 100 文字になるまで出力した. ここで新たに生成された文章を再び句点(“.”) で区切り, 水増し用データとして保存をした. この時, 100 文字で区切られた時に, 途中で生成した文章が終了した場合は, その文章を削除した.

以下に GPT-2 により生成した文章のみと, GPT-2 により生成した文章を元の文章に加えた場合の 5 クラス分類の分類精度を示す.

表 5 GPT-2 の分類精度

	分類モデル	分類精度
Original	BiLSTM-Attention	0.625
	BERT	0.500
GPT-2	BiLSTM-Attention	0.516
	BERT	0.510
Original + GPT-2	BiLSTM-Attention	<b>0.708</b>
	BERT	<b>0.789</b>

表から, GPT-2 によって生成した文章のみで分類した結果, BiLSTM-Attention は少し分類精度が悪くなったが, BERT は少し向上されることが確認できた. また, 生成した文章を元の文章に追加した場合, 両方のモデルとも精度が大幅に向上された.



### 6-3. 繰り返し学習

また、図 10 のように予測値と実測値の差が大きくなる箇所が複数見られるのは、5 クラス分類をする際の誤ったラベル付けなどが考えられ、このような予測ミスが減らすために行った繰り返し学習の結果を以下に示す。また、使用したモデルはいずれも GPT-2 で文章数の水増しを行ったものとする。

表 6 繰り返し学習による予測グラフの増減の正解率

初期モデル	通常モデル	分類精度	増減の正解率
BERT	BERT	<b>0.867</b>	0.635
BERT	BiLSTM-Attention	0.828	0.692
BiLSTM-Attention	BERT	0.855	<b>0.731</b>
BiLSTM-Attention	BiLSTM-Attention	0.805	0.692

表 7 繰り返し学習による予測グラフの評価指標

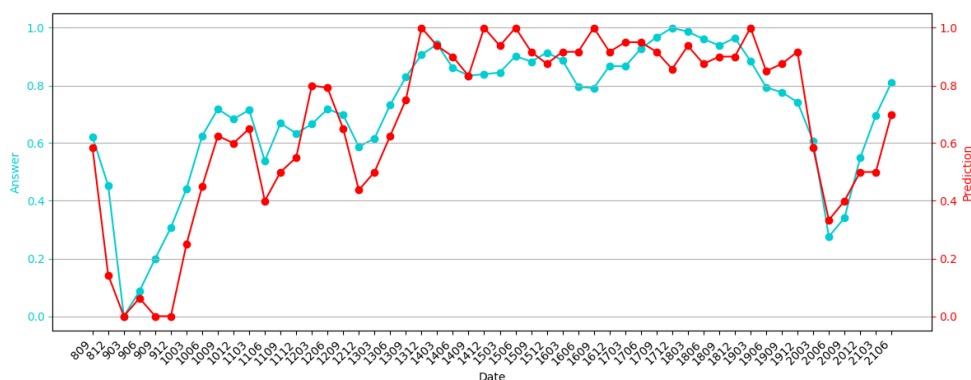
初期モデル	通常モデル	RMSE	MAE	AIC	R <sup>2</sup>
BERT	BERT	0.121	0.096	-98.971	0.724
BERT	BiLSTM-Attention	0.120	0.095	-81.78	0.729
BiLSTM-Attention	BERT	<b>0.114</b>	<b>0.094</b>	<b>-103.954</b>	<b>0.956</b>
BiLSTM-Attention	BiLSTM-Attention	0.189	0.127	-35.303	0.334

表から、BiLSTM-Attention と BERT を使用して学習を繰り返した結果、どの組み合わせでも分類精度が向上されることが確認でき、この中でも、BERT を初期モデルとして、その後も BERT で学習を繰り返した結果、分類精度が一番高くなることが確認できた。

また、BiLSTM-Attention を初期モデルとして、その後は BERT で学習を繰り返した結果、BERT を初期モデルとした場合よりも分類精度が少し劣るもの

の、4 つの指標とグラフの増減の正解率で他のモデルよりも上回ることが確認できた。

以下に BiLSTM-Attention を初期モデルとし、BERT を通常モデルとした際の景気動向の予測結果を示す。



## 7. 考察

### 7-1. 深層学習手法の精度に対する考察

まず GPT-2 を用いて文章数を水増ししない状態で、深層学習手法による景気文章の予測を行った際の評価をする。表 2 より、文章数 760 文の景気文章を分類した結果 BiLSTM-Attention における分類が一番良い精度となったが、他の深層学習手法と大きく差は見られなかった。実際、機械学習手法のサポートベクターマシンで同様の実装をした場合、分類精度が 0.537 となり、TextCNN には精度は少し及ばないものの、機械学習手法でも BERT より精度が高くなった。したがって、少量の文章を対象とした 5 クラス分類においては、どの手法にもおいては大きな差が出ることはなく、BiLSTM-Attention が他の手法よりも少し差が出たのは、Attention を加えることで文章の文脈情報をより理解することができたからと考えられる。

また、BERT モデルが少量の文章を対象とした分類の精度が低くなった際の評価をする。ここで、本研究で使用した BERT の事前学習済みモデルの語彙数は 32768 となり、あらかじめ 32768 の語彙を用いて、各入力文の関連性の理解などの言語パターンが学習されている。それに対して、BiLSTM-Attention を用いた学習時に使用した語彙数は、常に文章全体の語彙数と等しくなる。本研究における文章全体の語彙数は 796、GPT-2 を使用して文章を水増しした際の文章全体の語彙数は 1470 となる。したがって、BERT モデルで分類する際、語彙数が 796 の文章群に対して、語彙数が 32768 の学習済みモデルで分類を行うため、使用していない語彙の方が非常に多くなり、BERT モデルにおける分類精度が悪くなったと考えられる。

## 7-2. GPT-2 による文章生成に対する考察

GPT-2 を用いて文章数を水増しした際の、深層学習手法による景気文章の予測の評価をする。表 5 より、分類精度が大きく向上されたのは、生成した文章群が元の文章と同様の特徴を持っているからと考えられる。一般的に、深層学習モデルを使用した文章を対象とする予測には、大量の文章データと、その文章群が同様の特徴を持っている必要がある。本研究においては、大量の景気に対する文章が必要であり、GPT-2 により生成した文章群が景気に対する文章ではない場合、精度の向上に期待ができない。したがって、本研究では分類精度が大きく向上されたため、生成した文章は妥当であると言える。

また、通常の文章と GPT-2 を用いて生成した文章との分類精度で BERT のみ精度が向上されたのは、本研究で使った GPT-2 の事前学習済みモデル[16]の語彙数が 32000 であり、生成した文章により文章全体の語彙が増加し、BERT で学習する際に使用する語彙が増えたため、あらかじめ学習された言語パターンから文脈理解が容易になり分類精度が向上されたと考えられる。しかし、GPT-2 で文章数の水増しをした際の、予測グラフの増減の正解率は以下の表 8 に示す通り、BERT の方が劣ることが確認できる。

表 8 GPT-2 による予測グラフの増減の正解率

分類モデル	増減の正解率
BERT + GPT-2	0.673
BiLSTM-Attention + GPT-2	<b>0.692</b>

これは、BERT は学習する際の語彙数が 32768 個と固定されているが、BiLSTM-Attention で学習する際の語彙数は、常に文章全体の語彙数と等しく

なるため, BiLSTM-Attention は, 文章中にある語彙だけで言語パターンを学習することで, 景気文章の特徴のみを持つ言語パターンを理解し, より出力するグラフの正確な予測ができると考えられる.

### 7-3. 繰り返し学習に対する考察

繰り返し学習を行った際の深層学習手法による景気文章の予測の評価をする。表 6 で、初期モデルと繰り返しを行う通常モデルの両方を BERT + GPT-2 で行うことにより分類精度が本研究で一番高くなったのは、表 5 の GPT-2 を用いたモデルにおいて、BERT が BiLSTM-Attention よりも分類精度が高いためだと考えられる。

また、初期モデルが BiLSTM-Attention で、繰り返しを行う通常モデルを BERT にした際は、分類精度は少し劣るが、景気動向の予測グラフの 4 つの指標とグラフの増減の正解率で他のモデルよりも上回ることが確認できたのは、前述の「7-2. GPT-2 による文章生成に対する考察」で述べた通り、BERT はあらかじめ大量の語彙で学習された言語パターンを使用して学習を行うが、BiLSTM-Attention は景気文章の特徴のみを持つ言語パターンを学習する。これにより、初期に BiLSTM-Attention で学習を行うことで、景気文章に対する言語パターンを学習したベースモデルを作成し、その後、あらかじめ大量の語彙で学習された言語パターンを持つ BERT で学習を繰り返すことで、ベースモデルをより柔軟にし、正解との差が少ないグラフ予測をすることができたのだと考えられる。

## 8. 結論

少量データセットにおける予測を，日銀短観の調査文章を対象に深層学習手法を用いて予測をする手法の提案を行った．関連研究では内閣府の景気ウォッチャー調査などの大量データセットを対象に景気分析を行っているものが多かったが，本研究では，少量データセットという点に注目した景気動向の予測を行った．結果は汎用言語モデルである BERT と，BiLSTM-Attention の両方を使用することで柔軟な予測をすることができた．ただ，予測には正解との多少の差が出てしまうという課題点が見られた．この課題点は，様々な企業が掲載している予測文章における，景気状況に関連しない部分の排除をより緻密に行うことで，より正確な予測を行うことができると考えられる．

## 参考文献

- [1]. MUFG 三菱 UFJ リサーチ&コンサルティング, 「グラフで見る景気予報(1月)」  
URL: [https://www.m1urc.jp/report/economy/overall/graphical\\_japan/gr\\_2201/](https://www.m1urc.jp/report/economy/overall/graphical_japan/gr_2201/)
- [2]. 日本銀行, 「経済・物価情勢の展望(展望レポート)」  
URL: <https://www.boj.or.jp/mopo/outlook/gor2110b.pdf>  
閲覧日:2022 年 1 月 15 日
- [3]. 野村証券, 「証券用語解説集」  
URL: <https://www.nomura.co.jp/terms/japan/ke/keiki.html>  
閲覧日:2022 年 1 月 12 日
- [4]. NTT Communications, 「深層学習とは」  
URL: <https://www.ntt.com/bizon/glossary/j-s/deep-learning.html>  
閲覧日:2022 年 1 月 25 日
- [5]. 山本 裕樹, 松尾 豊: 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,(2016)
- [6]. 余野 京登, 和泉 潔: 金融レポート, およびマクロ経済指数によるリアルタイム日銀センチメントの予測, 第 31 回人工知能学会全国大会,(2017)
- [7]. 宮澤 和貴, 井上 輝也, 長井 隆行: GPT-2 を用いた人の動作予測, 第 35 回人工知能学会全国大会,(2021)
- [8]. 西 良浩, 菅 愛子, 高橋 大志: GPT-2 を用いた回答文生成ならびに LSTM による景気変動予測のテキスト分析, 第 13 回経営課題を AI に!ビジネスインフォマティクス研究会,(2019)



- [9]. 生田 祐介, 関 和広, 松林 洋一: テキストデータを利用した新しい景況感指標の開発と応用(下) 応用編: 深層学習を利用したテキスト分析, APIR Discussion Paper Series No.48,(2021)
- [10]. 日本銀行, 「短観」  
URL: <https://www.boj.or.jp/statistics/tk/index.htm/>  
閲覧日:2022 年 1 月 12 日
- [11]. 大和総研, 「2021 年 9 月日銀短観予想 製造業の急速な業況判断 DI の改善は一服へ」  
URL:  
[https://www.dir.co.jp/report/research/economics/japan/20210915\\_022524.html](https://www.dir.co.jp/report/research/economics/japan/20210915_022524.html)  
閲覧日:2022 年 1 月 12 日
- [12]. 斎藤 康毅, 「ゼロから作る Deep Learning 2 -自然言語処理編-」株式会社オライリー・ジャパン, (2018), p.223 - 244
- [13]. Dzmitry Bahdanau, Kyunghyun Cho, Yoshua Bengio: Neural Machine Translation by Jointly Learning to Align and Translate,(2014)
- [14]. 近江 崇宏, 金田 健太郎, 森長 誠, 江間 見亜利, 「BERT による自然言語処理入門 -Transformers を使った実践プログラミング-」株式会社オーム社, (2021), p.24 - 25, p30 - 41
- [15]. 東北大学乾研究室, 「Pretrained Japanese BERT models」  
URL: <https://github.com/cl-tohoku/bert-japanese>  
閲覧日:2022 年 1 月 14 日
- [16]. HuggingFace, 「GPT-2 small Japanese model」  
URL: <https://huggingface.co/colorfulscop/gpt2-small-ja>

閲覧日:2022 年 1 月 14 日

## 謝辞

本研究を進めるにあたり，熱心なご指導を頂きました法政大学応用情報工学科藤井章博教授に感謝いたします。また，様々な提案をしていただき，研究活動を支えてくださった藤井研究室の皆様にも感謝いたします。加えて，本研究に取り組むにあたり，法政大学入学，在学に多大なる協力をして頂いた両親にも感謝いたします。