

AIが予測する日本経済の先行き

敵対的生成ネットワークによる景気動向指数の短期予測

みずほ総合研究所
調査本部
03-3591-1298

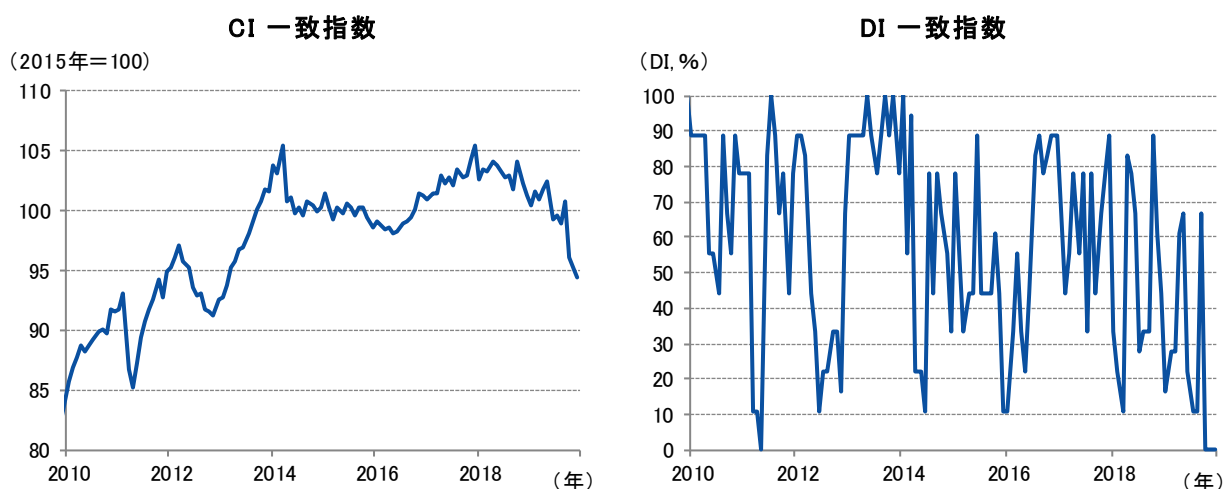
- 人工知能（AI）の一種である深層学習を用い、日本経済の動向を包括的に表す景気動向指数について、先行き3カ月間の短期予測を精度高く実施する試みを行った。
- 深層学習の4つの手法を比較した結果、主に画像生成などで用いられる敵対的生成ネットワークを改良したForGANモデルが、最も高い予測精度を示した。
- 現在利用可能なデータでは新型コロナウイルスの影響を考慮できないが、それでもモデルは2020年1～3月の更なる景気動向指数悪化を予測し、日本経済の足腰が弱まっていることが示唆される。

1. 昨年末に弱含んだ日本経済

2019年10～12月期の実質GDPは、前期比▲1.8%と5四半期ぶりのマイナス成長となった。落ち込み幅は、前回の消費増税直後（2014年4～6月期、前期比▲1.9%）と同等の大きさである。2019年10月の消費増税の影響を受け、GDPの半分以上を占める個人消費が大幅に減少したほか、民間企業の設備投資も3四半期ぶりのマイナスとなり、日本経済が2019年末に弱含んでいたことを示している。

今年に入って、新型コロナウイルスの影響により内外経済に多大な下押し圧力が加わっていることを踏まえると、2020年1～3月期も実質GDPのマイナス成長が続くのはほぼ確実である。今後は、景気が悪化する深さや速さを、より早いタイミングで正確に把握することが重要となる。

図表 1 景気動向指数



(注) 2020年3月6日に公表された2010年1月～2019年12月の値。
(出所) 内閣府「景気動向指数」より、みずほ総合研究所作成

そこで、みずほ総合研究所株式会社とみずほ情報総研株式会社は共同で、人工知能（AI）の一種である深層学習（ディープラーニング）を用い、日本経済の短期的な予測を精度高く実施する試みを行った¹。具体的には、深層学習の手法の一つである敵対的生成ネットワーク（Generative Adversarial Network、GAN）を応用したモデルを構築し、日本経済の動向を包括的に把握する指標である景気動向指数について、先行き3カ月間の動きを予測した。以下、その方法と結果を解説し、足元の日本経済に対する示唆について考察する。

なお、後段で詳述する通り、現在利用可能なデータには新型コロナウイルスの影響がまだ現れておらず、本稿で生成した2020年1～3月の予測値は「新型コロナウイルスの影響がなかったとしたら」という仮定の状況に基づくものである。今後、新たなデータの公表に伴い、新型コロナウイルスの影響を加味した予測ができるようになると考えられる。

2. 景気動向指数とは

景気動向指数は、生産、在庫、投資、雇用、消費、金融といった様々な経済部門の個別統計を合成することで、景気の現状把握や将来予測を目的として作成される指標であり、内閣府から毎月公表されている。

景気動向指数には、個別統計の動きを合成して作成するコンポジット・インデックス（CI）と、個

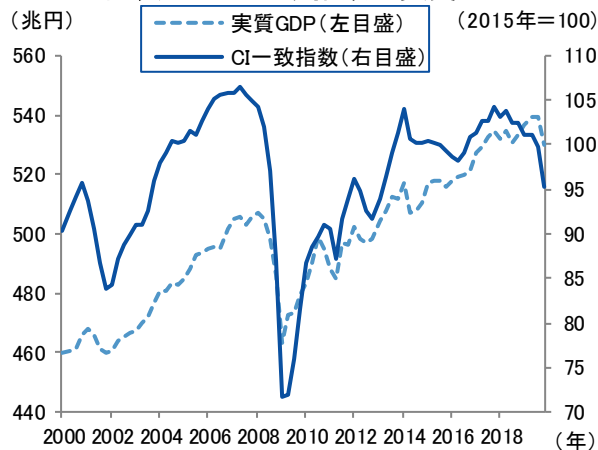
図表 2 景気動向指数の個別統計

先行指数	一致指数	遅行指数
1 在庫率指数(最終需要財)	1 生産指数(鉱工業)	1 第3次産業活動指数(対事業所サービス業)
2 在庫率指数(鉱工業用生産財)	2 出荷指数(鉱工業用生産財)	2 常用雇用指数(調査産業計、前年同月比)
3 新規求人数(除く学卒)	3 出荷指数(耐久消費財)	3 実質法人企業設備投資(全産業)
4 実質機械受注(製造業)	4 所定外労働時間指数(調査産業計)	4 家計消費支出(勤労者世帯、名目、前年同月比)
5 新設住宅着工床面積	5 出荷指数(投資財、除く輸送機械)	5 法人税収入
6 消費者態度指数	6 商業販売額(小売業、前年同月比)	6 完全失業率
7 日経商品指数(42種総合)	7 商業販売額(卸売業、前年同月比)	7 きまって支給する給与(製造業、名目)
8 マネーストック(M2、前年同月比)	8 営業利益(全産業)	8 消費者物価指数(生鮮食品を除く総合、前年同月比)
9 東証株価指数	9 有効求人倍率(除く学卒)	9 在庫指数(最終需要財)
10 投資環境指数(製造業)		
11 中小企業売上げ見通しDI		

(注) 2020年3月6日時点。

(出所) 内閣府「景気動向指数」より、みずほ総合研究所作成

図表 3 CI一致指数と実質GDP



(注) CI一致指数は2020年3月6日に公表された2000年1月～2019年12月の値。

(出所) 内閣府「景気動向指数」より、みずほ総合研究所作成

別統計のうち改善している割合を算出して作成するディフュージョン・インデックス（DI）がある。CIは景気変動の大きさやテンポをみるために、DIは景気の各経済部門への波及度合いをみるために用いられる（図表1）。

また、CIとDIにはそれぞれ、景気より先に動く先行指数、ほぼ同時に動く一致指数、遅れて動く遅行指数がある。各指数の動きが異なるのは、作成に用いる個別統計が異なるためだ（図表2）。先行指数は、在庫、新規求人、受注、住宅着工など、主に需給の変動を表す統計から作成される。一致指数は、生産、出荷、労働時間、商業販売といった、主に生産の調整を表す統計がもととなっている。遅行指数の個別統計は、雇用者数や失業率、設備投資、税金など、主に生産能力の調整や経済活動の結果を表している。すなわち、先行指数、一致指数、遅行指数は、経済活動を段階別に、順を追って捉えているといえる。

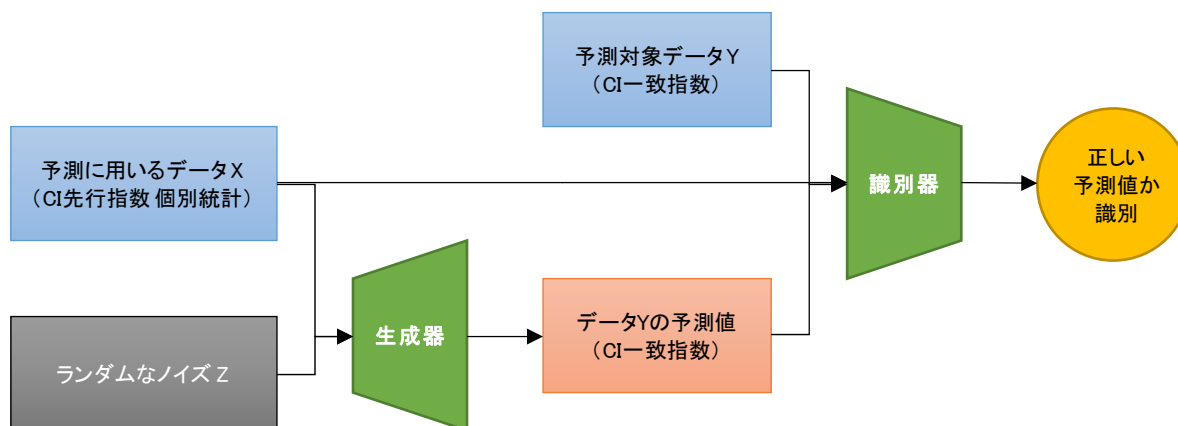
このように、景気動向指数には、作成方法の違い（CI、DI）と、作成に用いる個別統計の違い（先行指数、一致指数、遅行指数）により、合計6種類の系列が存在する。本稿で予測対象としたのは、このうちのCI一致指数である。CI一致指数は、景気の量的な変化を表し、かつ変化のタイミングが景気と一致するため、実質GDPと概ね連動した動きとなる（図表3）。したがって、景気動向指数のCI一致指数を精度高く予測することができれば、日本経済の動向を早いタイミングで把握できるようになるだろう。

3. 深層学習でベンチマークを上回る予測精度を達成

CI一致指数を精度高く予測するために、本稿では深層学習を活用した4つの手法、すなわち、リカレント・ニューラル・ネットワーク（RNN）、長・短期記憶（LSTM）、畳み込みニューラルネットワーク（CNN）、時系列GAN（ForGAN）について、予測精度の比較検討を実施した。その結果、最も高い予測精度を達成したForGANモデルを採用した²。なお、本稿における深層学習手法の実装と予測精度の検証は、みずほ情報総研が実施した。

ForGANモデルの基礎となっている敵対的生成ネットワーク（GAN）は、2014年に深層学習の研究者で

図表 4 ForGAN モデルの模式図



（出所）みずほ情報総研、みずほ総合研究所作成

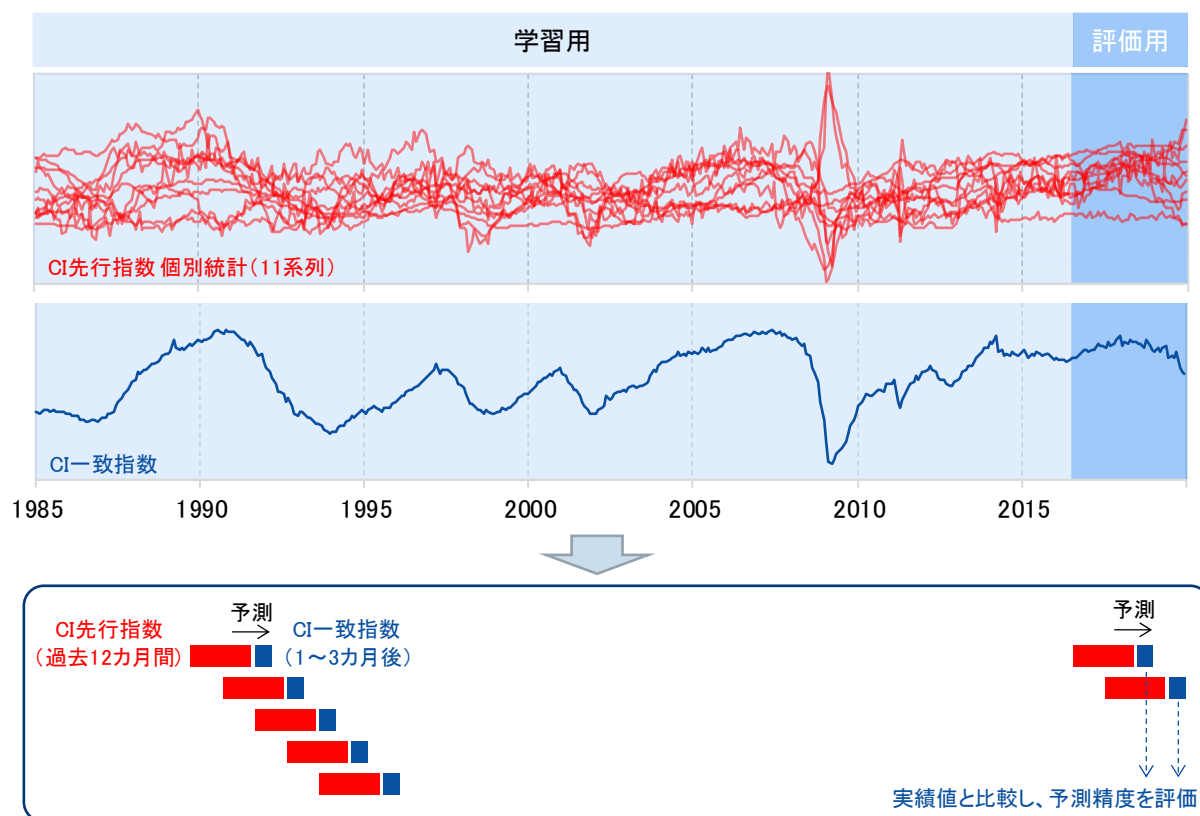
あるイアン・グッドフェローによって開発された手法である。GANは主に画像生成に用いられるが、現在では多くのバリエーションが誕生し、様々な用途に活用されている。以下、本稿で構築したForGANの仕組みについて、簡単に説明しよう。

ForGANでは、生成器（ジェネレーター）と識別器（ディスクリミネーター）という2つのニューラルネットワークを用いて予測を行う（図表4）。まず、予測に用いるデータ（X）とランダムなノイズ（Z）を生成器に与え、予測対象データ（Y）の予測値を生成する。次に、生成した予測値と、予測対象データ（Y）を識別器に与え、データを正しく予測できているか識別する。このように、生成と識別という2つのニューラルネットワークを競争させるように学習を進めていく（これが「敵対的」と呼ばれる所以である）。すると、次第に生成器は元のデータに近い予測値を生成することができるようになっていく。最終的に、識別器が元のデータと予測値を識別できなくなるほど、予測精度が高くなれば、学習が完了する。こうしたモデルの仕組みを踏まえ、実際にモデルを構築する手順をみてみよう。

モデルを構築するうえで最初に重要なポイントは、予測対象（図表4のY）であるCI一致指数を予測するために、どのようなデータ（図表4のX）を使うかである。本稿では、データXとして、CI先行指数の個別統計（11系列）の過去12カ月間の値を用い、1～3カ月先のCI一致指数を予測することとした。CI先行指数の個別統計11系列を用いたのは、経済活動の早い段階の動きを、幅広くモデルに取り込むことができると考えたからだ。

次に、本稿執筆時点で利用可能な1985年1月～2019年12月のデータ（2020年3月6日公表）を、9割（1985

図表 5 データの分割と予測・評価方法の概念図



(出所) みずほ情報総研、みずほ総合研究所作成

年1月～2016年6月)の学習用データと、1割(2016年7月～2019年12月)の評価用データに分割した(図表5)。学習用データでモデルの学習を実施した後、評価用データを与えて予測値を計算し、精度を評価することで、モデルがまだ知らないデータに対しどの程度精度で予測ができるかを確認することができる³⁾。

なお、深層学習では一般的にデータ全体をランダムに分割し、学習用データと評価用データを生成する。しかし、本稿のような時系列分析では、データをランダムに分割すると、学習用データと評価用データが重複する期間が生じてしまい、評価用データを用いて求めた予測精度が未知のデータに対する予測精度とならず、予測精度を適切に評価することができない。したがって、本稿では、時系列上の一時点(2016年7月)を閾値として定め、それより前を学習用、以後を評価用としてデータを分割した。

このようにモデルの学習と予測精度の評価を行った結果を、図表6にまとめた。記載した値は、各モデルの予測値と実績値の乖離度合いを表す平均二乗誤差(Mean Squared Error, MSE)であり、MSEの値が小さいほど、予測精度が高いことを意味する。なお、比較のためのベンチマークとして、CI一致指数自身の過去の値を用いて予測した結果(自己回帰)と、CI先行指数の個別統計を用いて伝統的な線形モデルで予測した結果(重回帰)もあわせて掲載した。

図表6を見ると、4つの深層学習モデルのうち、RNNとLSTMはベンチマークの自己回帰に比べ予測精度の向上がみられなかったが、一方でCNNとForGANはベンチマーク対比で予測精度が大きく向上した。とりわけ、ForGANのMSEが0.7～1.2台と小さく、ベンチマークモデルと深層学習モデル全体で最良の予測精度を達成した。

一般的に、時系列データの予測を行うと、先の予測になるほど精度が悪化する(MSEが大きくなる)傾向がある。ForGANモデルでは、1カ月先予測に比べ、2・3カ月先予測のMSEが大きく、2・3カ月先予測の精度が1カ月先予測に比べやや悪いことが確認できる。しかし、それでもForGANモデルの2・3カ月先予測のMSEは1.2台にとどまり、他のモデルの1カ月先予測のMSE(1.3～)より小さい。これは、ForGANが先々の予測になっても相対的に高い精度を維持していることを示す結果である。

図表7は、CI一致指数の実績値と、ForGANモデルを用いて生成した1カ月先予測値、2カ月先予測値、3カ月先予測値をそれぞれ図示したものである。1カ月先予測値は、実績値の動きをほぼ捉えているこ

図表 6 モデルの予測精度

		1カ月先予測	2カ月先予測	3カ月先予測
ベンチ マーク	自己回帰	1.69	2.49	3.59
	重回帰	6.15	8.75	13.25
深層学習 モデル	リカレント・ニューラル・ネットワーク(RNN)	2.61	2.47	3.57
	長・短期記憶(LSTM)	2.72	4.54	4.80
	畳み込みニューラルネットワーク(CNN)	1.30	1.69	1.85
	時系列GAN(ForGAN)	0.77	1.25	1.24

(注) 数値は平均二乗誤差(Mean Squared Error, MSE)であり、値が小さいほど精度が高いことを示す。

(出所) みずほ情報総研作成

とが確認できる。2カ月先予測値と3カ月先予測値は、部分的に実績値から乖離している時期もあるが、概ね実績値の傾向に沿って推移しているといえるだろう。

4. 予測モデルは2020年1～3月の更なる景気動向指数悪化を示唆

最後に、本稿執筆時点でCI一致指数と先行指数が入手可能な2019年12月までのデータ（2020年3月6日公表）すべてを用いてForGANモデルの再学習を行い、2020年1～3月のCI一致指数を予測した（図表8）。ForGANモデルによる2020年1～3月の予測値は、1月に低下した後、2月にはやや反発するが、3月には再び低下している。四半期ごとの平均値をとると、2019年7～9月期実績値が99.7、同10～12月実績値が95.2、そして2020年1～3月期予測値が93.8となり、CI一致指数が今年に入って更に悪化することが示唆される。

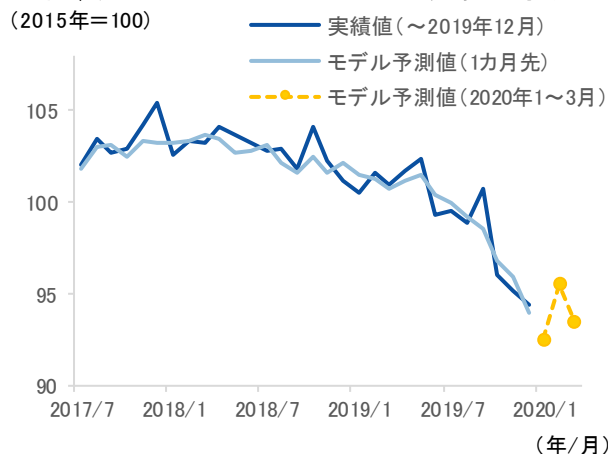
なお、2020年1～3月期の予測値を計算するために用いた再学習データ（～2019年12月）には、年明け以降の新型コロナウイルスによる影響がまだ現れていない。したがって、ForGANモデルによる2020年1～3月の予測値は、「もし新型コロナウイルスによる経済への影響がなかったとしたら」という仮定

図表 7 CI一致指数実績とForGANモデル予測



(注) CI一致指数の実績は3月6日公表の値。
 (資料) 内閣府「景気動向指数」より、みずほ情報総研、みずほ総合研究所作成

図表 8 ForGANモデルによる先行き予測



(注) CI一致指数の実績は3月6日公表の値。
 (資料) 内閣府「景気動向指数」より、みずほ情報総研、みずほ総合研究所作成

の状況に基づくものである。

現実には、新型コロナウイルスの流行拡大防止のために国内で各種イベント・興行が自粛されているほか、欧米諸国をはじめとする入国制限や都市の封鎖により、世界的に人の移動が妨げられ、内外の経済活動に大幅な制限が課されている。こうした状況に鑑みれば、今後公表されるCI一致指数の実績値が、本稿のモデルによる予測値を大きく下回るのは間違いない。重要なのは、こうした新型コロナウイルスによる影響がなかったとしても、日本経済は2020年1～3月に一段と低迷していた可能性が高いということが、モデルから示唆される点である。それだけ、日本経済の足腰が弱まっていたということだ。

今後は、新型コロナウイルスによる影響を含む最新データをモデルが順次学習していくことで、その影響を考慮した予測値を生成することができるようになると考えられる。足元の日本経済の低迷が、どの程度の速さで、どこまで深くなるのか、そして、景気の底入れがいつ訪れるのかをいち早く知るうえで、本稿のモデルが活用できるだろう。

[参考文献]

- A. Koochali, P. Schichtel, A. Dengel and S. Ahmed, "Probabilistic Forecasting of Sensory Data With Generative Adversarial Networks - ForGAN," in IEEE Access, vol. 7, pp. 63868–63880, 2019.

1 本共同研究のメンバーは下記のとおりである。

みずほ情報総研株式会社

事業戦略部 AI Powerhouse

事業戦略部

事業戦略部

事業戦略部 AI Powerhouse

情報通信研究部 AI Powerhouse

情報通信研究部

みずほ総合研究所株式会社

経済調査部長 兼 高度デジタル情報解析室長

経済調査部 兼 高度デジタル情報解析室

経済調査部

経済調査部

高度デジタル情報解析室

市場調査部 兼 高度デジタル情報解析室

Chief AI Officer

参事役

調査役

調査役

主任研究員

コンサルタント

笠間俊夫

羽田謙一郎

大西史花

井筒雄介

小泉拓

加田匠

太田智之

主任エコノミスト

宮嶋貴之

主任エコノミスト

小寺信也

エコノミスト

矢澤広崇

主任研究員

服部直樹

エコノミスト

越山祐資

2 ForGAN モデルは、敵対的生成ネットワーク (GAN) の応用例の一つである Conditional GAN を用い、時系列データをジェネレーターとディスクリミネーターに付与する手法。詳細は、Koochali, et al. (2019) を参照されたい。なお、本稿で構築した ForGAN モデルは、Koochali, et al. (2019) を一部改良したものである。

3 評価用データは 2016 年 7 月～2019 年 12 月だが、CI 先行指数の個別統計 (11 系列) の過去 12 カ月間の値をもとに予測を行っていることから、予測精度の評価期間は、1 カ月先予測が 2017 年 7 月～2019 年 12 月、2 カ月先予測が 2017 年 8 月～2019 年 12 月、3 カ月先予測が 2017 年 9 月～2019 年 12 月、となる。

[共同執筆者]

みずほ総合研究所株式会社	高度デジタル情報解析室	主任研究員	服部直樹	naoki.hattori@mizuho-ri.co.jp
みずほ情報総研株式会社	事業戦略部 AI Powerhouse	調査役	井筒雄介	yuusuke.izutsu@mizuho-ir.co.jp

●当レポートは情報提供のみを目的として作成されたものであり、取引の勧誘を目的としたものではありません。本資料は、当社が信頼できると判断した各種データに基づき作成されておりますが、その正確性、確実性を保証するものではありません。本資料のご利用に際しては、ご自身の判断にてなされますようお願い申し上げます。また、本資料に記載された内容は予告なしに変更されることもあります。なお、当社は本情報を無償でのみ提供しております。当社からの無償の情報提供をお望みにならない場合には、配信停止を希望する旨をお知らせ願います。