

IMES DISCUSSION PAPER SERIES

ビッグデータと人工知能を用いた
ファイナンス研究の潮流

いずみ きよし
和泉 潔

Discussion Paper No. 2018-J-19

IMES

INSTITUTE FOR MONETARY AND ECONOMIC STUDIES

BANK OF JAPAN

日本銀行金融研究所

〒103-8660 東京都中央区日本橋本石町 2-1-1

日本銀行金融研究所が刊行している論文等はホームページからダウンロードできます。

<https://www.imes.boj.or.jp>

無断での転載・複製はご遠慮下さい。

備考：日本銀行金融研究所ディスカッション・ペーパー・シリーズは、金融研究所スタッフおよび外部研究者による研究成果をとりまとめたもので、学界、研究機関等、関連する方々から幅広くコメントを頂戴することを意図している。ただし、ディスカッション・ペーパーの内容や意見は、執筆者個人に属し、日本銀行あるいは金融研究所の公式見解を示すものではない。

ビッグデータと人工知能を用いたファイナンス研究の潮流

いずみ きよし
和泉 潔*

要 旨

近年、多くの分野においてビッグデータと人工知能技術の応用が進んでおり、ファイナンス分野もその例外ではない。本稿では、その最新事例を概括し、今後のさらなる発展の方向性や克服すべき課題について議論を行う。最初に、これまで定量分析が困難であった画像やテキストといった非構造化データおよび大規模データについて、機械学習を用いて分析し、金融実務や市場分析に取り入れた最新事例を紹介する。次に、ファイナンス分野への人工知能応用の課題として、問題設定能力・推定過程の透明性・他者の反応の推定の3点について論じる。最後に、これらの課題を克服するための手法の1つとして、複数のプログラム同士を人工市場において対戦させる、自己対戦型の学習を紹介する。ファイナンス分野で利用されている人工知能技術には、大きな期待が寄せられているものの、同技術は万能ではなく、現時点では、あくまで人間の能力を拡張するツールに過ぎない。今後は、特定の状況において有効な個々の人工知能技術を複数組み合わせることで、複雑な状況にも対応できるよう高度化を進めることが望まれる。

キーワード：ビッグデータ、人工知能、テキストマイニング、機械学習、
金融市場分析

JEL classification: C31、C32、C45、D83

* 東京大学大学院工学系研究科教授 (E-mail: izumi@sys.t.u-tokyo.ac.jp)

本稿は、2018年3月5日に日本銀行金融研究所が開催した「ビッグデータと人工知能を用いたファイナンス研究の展開」をテーマとするファイナンス・ワークショップにおける基調講演の内容をまとめたものである。ただし、本稿に示されている意見は、筆者個人に属し、日本銀行の公式見解を示すものではない。また、ありうべき誤りはすべて筆者個人に属する。

1. はじめに

近年、人工知能（artificial intelligence: AI）技術は第3次ブームを迎え、社会のさまざまな分野で応用されて話題となっている。人工知能技術とは、機械で人間と同じような知能を実現する技術であり、例えば、車両の自動走行や医療の自動診断、人間との対話システムは人工知能技術によりその能力を大幅に向上させた。少し前まで当分は機械には無理だと思われていた、将棋や囲碁等の知的なゲームの熟達したプレイや、絵画や音楽等の芸術作品の創造までも、人間並みか、場合によっては人間を上回るパフォーマンスでこなすことが可能となった。

こうした人工知能技術の多分野における応用の急速な進展の背景には、次の3つの要素がある。第1に、センサや通信・情報処理技術の発展や普及に伴い、いわゆる「ビッグデータ」と呼ばれる、今までよりも大規模で多種類のデータが獲得できるようになったことである。人工知能技術を実社会に応用するためには、大規模な実データの分析が欠かせない。第2に、これらの大規模データを分析して、データの中にある有用なパターンを見つけるための機械学習（特に深層学習）の技術が発展したことも、人工知能の発展の大きな要因である。機械学習は人工知能の基礎技術の1つであり、機械（コンピュータ）を使ってデータから新たな知識やルールを認識し獲得する技術である。近年は、深層学習（deep learning）等の発展した機械学習手法の登場により、画像や音声のデータ解析分野では、人間が機械にほとんど前提知識を与えなくとも、大規模データから自動的に有効なパターンを見つけ出せるようになってきた。第3に、大規模なデータを対象としてパターン発見のための高度な計算を高速に実行できる、大規模並列計算機技術の進展である。従来は現実的な時間内にパターンを発見できなかった大規模なデータからも、計算速度の向上によって、実務に耐える時間内でパターンを発見できるようになってきている。

このような技術的・学術的な発展を背景に、多くの分野においてビッグデータと人工知能技術の応用が進んでおり、ファイナンス分野もその例外ではない。従来、データや処理速度の制約から、経済活動は現在までの経済活動指標のみによって決まるものと想定し、先行きを予測することが多かった。しかし実際の経済活動は、経済活動指標に限らず、世の中のさまざまな出来事から影響を受けている。したがって、現在および過去に起きてきたあらゆる出来事について、より多く、より早く、そして正確に知ることができれば、将来の経済動向を他人よりも早く高い精度で予測できるはずである。近年の情報通信技術の進歩によって、そうした状況に近づきつつあると考えられるものの、現時点では、金融分野での応用には乗り越えなければならない課題がまだまだ多く存在する。

本稿では、さまざまなデータを金融市場で活用する技術が、今日および将来の金融市場の現場にどのような変化をもたらしつつあるのか、いくつかの具体的な事例を紹介し、その方向性と克服すべき課題を議論する。

2. ファイナンスへの人工知能応用の現状

人工知能技術の応用・発展の波は、ファイナンス分野にも押し寄せている。特に、資産運用の分野では、もともとクオンツ分析等で定量的なデータを統計的に解析する試みが普及しているため、人工知能技術の中でも機械学習の手法との相性が良く、同技術を用いて今までとは規模が格段に違うテラバイトやペタバイトの膨大な金融データを高速に自動解析することが可能となってきた。そうした数値データに加え、画像や音声、言語といった非構造化データにまで分析対象を拡張し市場分析に組み入れることも、データ取得技術や人工知能技術の高度化に伴って可能となってきた。実際に、そのような人工知能によるビッグデータ活用技術を利用して資産運用していることをアピールする金融機関も登場してきているほか、同技術により得られた情報を金融機関や個人投資家に提供する情報ベンダーも出現してきている。以下では、活用するデータの形式ごとに、そうしたビジネスや研究の具体的な事例を紹介する。

(1) 画像データの活用

センサ技術とデータ解析技術の進歩がもたらした恩恵として最初に挙げられるのが、大量の画像の分析が可能となった点である。ある種の地表観測用人工衛星が地表をとらえた画像では、地上の数十センチ四方の大きさまで観測できる解像度が得られる。複数のそうした人工衛星が撮影した画像を集めることによって、地球上のほぼ全域の画像を1日数回以上定期的に取得することができる。2014年の米国の規制緩和により、このような高解像度の衛星画像をさまざまな目的で一般企業が利用することが可能となった。例えば、グーグル社はデジタルグローブ社から購入した高解像度の衛星画像をグーグルマップの地図サービスに使っている。他にも災害復旧や都市計画、資源探索、航路探索、自然観測等の商用利用が進んでいる。

当然、金融市場の分析にこれらの衛星画像を用いるサービスも商用化されている。例えば、米国のオービタル・インサイト社 (Orbital Insight <https://orbitalinsight.com/>) は、衛星画像から特定地域の石油貯蔵量を推定し、石油市場の動向を予測するサービスを提供している。具体的には、複数の会社から購入してきた衛星画像から、機械学習による画像解析技術を用いて石油タンクとその浮

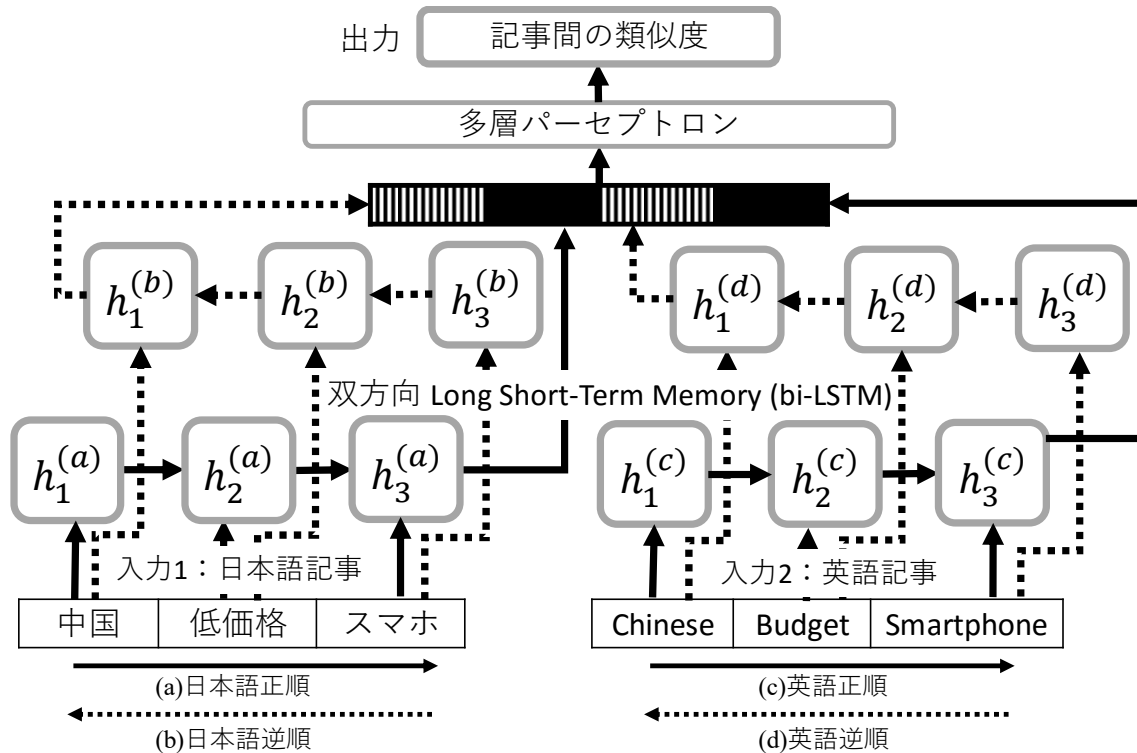
き屋根の影の変化を検出することで、タンクに貯蔵されている石油の量を推定している。これにより、正確な石油貯蔵量の統計値が入手困難な地域においても、信頼性の高い推定値を毎日得ることができ、投資会社やヘッジファンド等の顧客により市場予測に活用されている。また、同じ米国のテルアスラボ社 (TellusLabs <https://telluslabs.com/>) は、米国海洋大気庁からの気象データや、米国農務省からの季節と作物の成長情報のほかに、機械学習を用いて高解像度の衛星画像を分析することで特定地域の農業生産を予測し、先物市場分析のための情報を提供するサービスを行っている。

(2) テキスト・データの活用

画像以外の新たなタイプのデータでは、特にテキスト情報の解析による市場分析 (金融テキスト・マイニング) が 1990 年代後半から始まった (Nassirtoussi *et al.* [2014])。初期の金融テキスト・マイニングは、新聞記事やニュース記事を対象とした分析が多かったが、ソーシャル・メディアの普及に伴い、2000 年代からはツイッターやブログ等のテキスト・データを金融市場予測に用いる研究も増加してきた。例えば、ボレンらは、2008 年 2 月 28 日から 11 月 28 日の 9,853,498 個のツイッター・データを分析し、米国のダウ・ジョーンズ工業株価平均との関係性を調べた (Bollen, Mao, and Zeng [2011])。「楽しい」「悲しい」等のユーザの心理状態に関連する表現の出現頻度に着目して分析したところ、翌日の平均株価の騰落の方向性を 86.7%の精度で予測でき、テキスト・マイニングが株価動向の予想に有効であるとの実証分析結果を得た。

また、テキスト・マイニングを市場分析に役立てる試みとして、複数言語で書かれた金融記事を人工知能技術によって自動的に対応付ける自動マッピングという研究もある (Liu *et al.* [2017])。図 1 のように、英語の経済ニュース記事と日本語記事の類似度を機械学習の一種である多層パーセプトロンと双方向 Long Short-Term Memory (bi-LSTM) と呼ばれる手法により計算し、多言語テキスト間の類似性判断を行う。機械学習の入力は日本語記事と英語記事の単語系列であり、与えられた多言語記事間の類似度を 0 から 1 までの値で出力する。1,000 組の日英翻訳の経済ニュース記事でテストしたところ、68.5%の高い精度で翻訳の組を当てることができた。この技術を応用すれば将来的に、自分に馴染みのない言語で書かれた世界中の経済ニュースに対して、自分の知っている言語で書かれたよく似た記事を機械に提示させ、世界中のさまざまな経済ニュースの内容を素早く理解できるようになる。

図1 多言語金融記事間の自動マッピングの枠組み



備考：図中の $h_i^{(s)}$, $i = 1, 2, 3 \dots, s = a, b, c, d$ は bi-LSTM の中間層を意味する。 i は各記事に出現した単語の順番を表す。 $s = a$ は日本語記事を頭から（正順に）入力することを意味し、 b は日本語の逆順、 c は英語の正順、 d は英語の逆順の処理を示す。

現在、このような金融テキスト・マイニング技術を実際の資産運用に用いている運用会社のほか、中央銀行や企業の公表したテキストを定量化する民間サービスも数多く存在する。例えば、プラットル社 (Prattle <https://prattle.co/>) は、米国や日本をはじめとした各国の中央銀行の発行物をテキスト・マイニングし、そのセンチメント（テキスト内容のポジティブとネガティブの度合い）をリアルタイムに配信するサービスを行っている。具体的には、発行物ごと、要人の発言ごとに機械学習の手法でセンチメントを計算し、その移動平均をスコアとして算出している。また、インサイト 360 (Truevalue Labs <https://www.truevalue labs.com>) というサービスでは、企業に関するテキスト・データや数値データ等入手し、ESG (Environment, Social, Governance) 評価に関する 14 のトピックの指数を生成している。今までは、企業の ESG 評価は人が行っていたためバイアスが含まれていたが、大量のデータを基に評価することで客観性を担保することを目指している。これらに加え、ブルームバーグやトムソン・ロイター、日経 Quick 等の既存の情報配信会社も、ニュース記事を機械で解析しポジティブかネガティブかのスコアを付加したり、企業や公共機関の発表資料から自動的に要約を作成したりするなど、人工知能技術によるニュース解析を始めてい

るほか、さまざまなベンチャー企業もこの分野で立ち上がっている。

このように、非構造化データを含む大規模データを人工知能技術によって解析することで、これまで人手では処理しきれなかった情報を資産運用や市場分析に取り入れることを可能とするビジネスや研究が次々に登場してきており、この傾向は今後より一層強くなるであろう。

3. ファイナンスへの人工知能応用の課題と将来

2 節で述べたように、大規模データの獲得と解析技術の高度化によりファイナンス分野への人工知能技術の応用が進んできたが、同技術も万能ではなく、人間と同様の知的作業を行うためには技術的な課題も数多く残されている。

問題点を明らかにする出来事として、2013 年のツイッター・クラッシュ事件がある。2013 年 4 月 23 日午後、ダウ工業株 30 種平均指数が突然 140 ドル以上急落し、米国株式市場が大混乱となった（日本経済新聞 [2013]）。原因は、通信社のツイッター・アカウントから流された偽のテロ情報だとみなされている。偽ニュースに含まれていた「ホワイトハウスで爆発」「大統領が負傷」といったキーワードに、自動売買プログラムが反応してしまったのである。同プログラムに採用されていた金融テキスト・マイニング技術は、ニュース等のテキスト情報をコンピュータによって自動処理し市場分析するもので、1990 年代から金融市場で普及してきたが、人間が予め与えた特定のキーワードがニュースに現れた場合に決められた取引を行うという、単純なキーワード・マッチ型の技術であったため、上述の偽ニュースに騙されてしまったのである。

人工知能技術、特に機械学習は、過去の大量のデータに照らし合わせて、現在の状態と似た過去の状況を見つけ出し、将来の動向を推定することは得意である。しかし、大所高所にたって自分の推定結果が有効かどうかを判断することは苦手である。特定の情報による推定だけでなく、一般常識や多様な情報との一貫性の確認、他の人たちや世の中の反応のチェック等もできれば、上述の偽ニュースにだまされることはなかったかもしれない。以下では、人工知能技術の信頼できる適用のために必要な課題を論じる。

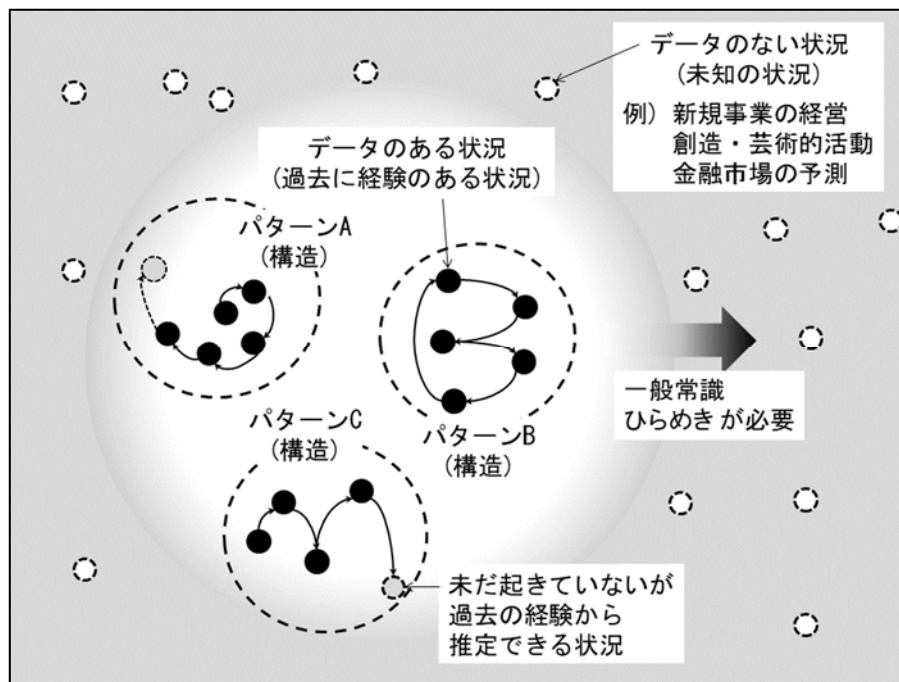
(1) 問題設定能力

車両の自動走行や医療の自動診断、将棋や囲碁のプレイ等、少し前まで当分は機械には無理だと思われていた分野でも、機械が人間並みか、場合によっては人間を上回るパフォーマンスでこなすことが可能となった。しかし、これら

全てのタスクにおいて、最初の問題設定は人間から与えられており、機械は決められた環境と判断材料の中で、その問題設定に特化して最良と思われるパターンを見つけ出しているのみである。例えば将棋や囲碁の学習では、行動の選択肢（駒や石の動かし方）や評価方法（王を取られたら負け、相手より陣地を多く囲った方が勝ち）、参照すべきデータ（過去の棋譜データ）等の問題設定は、予め人間が決定している。機械はどんな行動があり得るのか、この評価自体が間違っていないか、どのデータを参照すべきかという問題設定自体にかかわる根本的な問題には悩まずにすむ。

それに対して金融市場においては、市場参加者が市場の指標以外にも世の中一般にある情報を参照して行動し、その行動が金融市場に影響を与えており、潜在的には世の中の全ての事柄が関わり得る。例えば、太陽の黒点の数に関するデータも金融市場分析に使うべきか等、参照するデータと参照しないデータの線引きを判断していかなければならない。また、取引戦略は常に進化しているので、行動の選択肢も常に新しいものが出てくる。場合によっては、自分自身で新たな取引戦略を開発することもあり得る。さらに、自分が選択した行動の評価も、その後の金融市場の変動によって判断しなければならないこともある。例えば、その時々を経済的な状況によって、リターンを重視すべきか、リスクを重視すべきか、短期的な利得を重視すべきか、長期的な利得を重視すべきか等を自分で判断しなければならない。時には、図 2 に示すような過去データに無い未知の状況も想定して問題設定をしなければならない。

図2 社会現象の問題設定の範囲とデータ解析の有効範囲



資料：和泉・齋藤・山田 [2017]

このような、問題設定の根本的な課題や、あらゆる出来事にかかわる一般的な問題を解くことは、現時点では人間の方が機械よりもはるかに優れている。おそらく人間は、それまでのさまざまな社会的な実体験に基づいた「一般常識」や「ひらめき」を使って、こうした問題を解いているが、機械にはまだそれができないためである。したがって、金融市場分析での機械と人間の当面の役割分担は、人間がまず関係のありそうなデータの範囲や目標を示し、そこから機械学習によるデータ解析を用いて有効なパターンの候補を機械に挙げてもらい、機械が提示した候補をどう評価して実際の投資に使うのかを人間が判断するというものになるであろう。さらに、自分が持っている過去データに無かったような新しいイベントや急激な変化が発生した場合の大局的な判断には、依然として人間の常識や直観による判断が求められるであろう。

(2) 推定過程の透明性

適切な問題設定ができて、機械が何らかの予測結果を出力したとしたら、その予測結果が信頼できるものかどうか評価を行わなければならない。たとえ過去データによくフィットしていても、将来の運用で、ある日突然予測精度が低下することが起きる可能性があるため、過去データによるバックテストだけでは不十分である。また、複数の手法による市場予測を行って、過去データのバ

ックテストで同等の精度を示した場合に、どちらの手法を使えばよいかという判断も必要になることがある。そうした判断を人間が下すにあたっては、機械がどのような推論過程を経てその結果に至ったのかが、人間に理解可能である必要がある。

線形回帰等の伝統的な統計解析ならば、データのどの要素がどう作用して、このような予測結果となったのかという予測の根拠やプロセスが、人間にも理解可能である。ある要素のこれからの動向が不明確な時期であれば、その要素が作用して出た予測結果の信頼性は低くなるであろう。逆に、ある要素の動向が確実にわかっている場合には、その要素を重要視した予測結果の信頼性は高い。予測のプロセスが人間にも理解可能な場合は、このような判断が可能である。しかし、深層学習等の高度な手法によって学習した結果の中身は、複雑すぎて人間には理解できない、ブラック・ボックスになってしまうことが多い。外れるにしろ当たるにしろ、その理由がわからない手法に、実際の資産運用を全て任せることは無理であろう。

機械学習によるデータ解析の結果を、人間に理解可能なようにするための技術的な試みは、いくつか行われている。例えば、ある機械学習手法によって得られた結果を、別の機械学習手法により人間に理解しやすいような形に変換する研究がある (Raghu *et al.* [2017])。他にも、学習結果が人間に理解しやすく、しかも予測精度ができるだけ落ちないように、機械学習手法に制限を加える方法も研究されている (和泉ほか [2017])。また、与えるデータの種類を変えながら、別々の機械学習手法を使い、それぞれの結果を総合して予測するという手法も研究されている (Fujimaki, Sogawa, and Morinaga [2011])。例えば金融市場では、マクロ経済指標等の長期的な市場動向に影響を与えそうなデータを用いた学習結果と、日中価格等の短期的な動向に影響を与えそうなデータによる学習結果を合わせて、ある状況では長期要因による予測結果を重要視し、またある状況では短期要因による予測結果を重要視するという、切り替えを学習するメタなレベルの解析も行うことが考えられる。このような複数のデータ分析の統合によって、その時々相場を支配している式を推定し、それぞれの学習結果の信頼性を判断していくことも、これからの挑戦として必要である。

(3) 他者の反応の推定

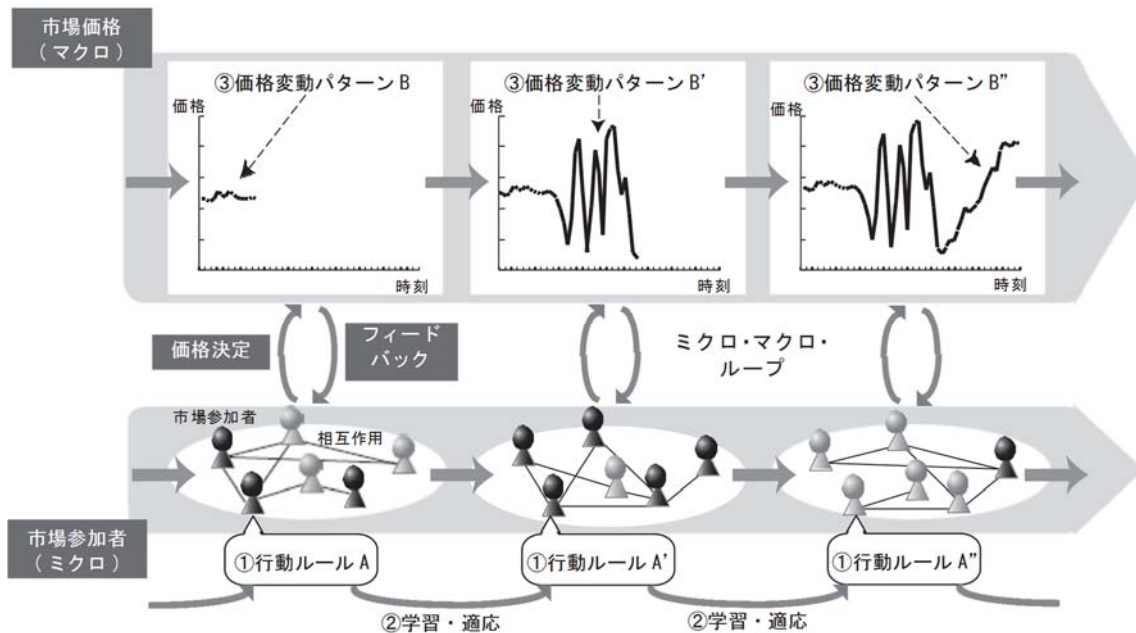
自らの投資行動を含めた、世の中のさまざまな出来事に対する他者の反応がすべて推定できれば、正確に市場を予測することが可能となる。しかしそれには、以下に挙げる 3 つの問題があり、機械学習を用いる場合だけでなく、人手で分析する場合も含めて難しい。

一番難しい問題は、金融市場の構成要素である市場参加者個人の行動がすでに複雑であり、物理現象の粒子や流体のような基本方程式が無いことである。金融市場でのトレードや商品の購入等の一般の経済社会的行動について、全ての人々の行動を普遍的に説明し、予測できる行動原理は無い。もしそのような普遍的な行動原理が存在したとしても、今のところは解明されていない。これが物理学や化学等の自然科学的現象と、人間行動が絡む経済社会現象との一番の大きな違いである。自然科学的現象には、その現象を構成する要素の挙動を支配している基本方程式が解明されている場合が多い。基本方程式を用いれば、構成要素の挙動は安定して正確に説明できる。またその要素が構成する現象全体の振舞いも、基本方程式の合成によりある程度予測できる。そのため、分析したい自然科学的現象について、たとえ過去に同じ状況や条件が無かったとしても、その状況でその現象がどのように振舞う可能性があるかを、既存のデータを基にして推定することができる。つまり、未知の状況に対して外挿予測ができるのである。これに対して金融市場では、未だ起きていない状況を持つ特徴を、過去のある程度類似した状況から基本方程式によって予想することが困難であり、全て未知の状況になってしまう。しかも、市場参加者は知能を持っており、状況を見ながら行動ルールを自分で変えてしまう。

2つ目の問題は、構成要素である市場参加者の挙動が均質でないことである。たとえ市場参加者の行動原理が単純な状況になっていたとしても、各個人が持っている取引や予測のルールはそれぞれ異なっているであろう。そのため、金融市場全体の振舞いを決定する条件も膨大な数になる。さまざまな状況に関してデータを大規模に集めようとしても、全ての状況をデータで網羅することは不可能である。そのため、過去データに無い未知の状況が数多く存在することになる。

3つ目の問題は、マイクロ・マクロ・ループの存在である（図3）。たとえ市場参加者個人（マイクロ）の行動原理が単純で、市場全体で均質だとしても、個人の行動が他の個人の行動と相互に影響していると、金融市場全体（マクロ）の振舞いは複雑で予想できないものになり得る。そのため、金融市場の状況は多様になり、過去データの解析だけではカバーできない、新たな状況が将来発生する可能性は常にある。

図3 金融市場でのマイクロ・マクロ・ループ



資料：和泉・齋藤・山田 [2017]

金融市場の予測、あるいは変動メカニズムの理解に関するこれらの根本的な課題への取組みの1つとして、複数の機械学習プログラム同士で取引を行い、金融資産価格が変動するプロセスをシミュレートする研究が行われている（和泉 [2012]）。こうしたプログラム同士の対戦（自己対戦）による学習は、過去データにない新たな状況における他者の反応も知ることができる有用な方法であり、囲碁の対戦プログラムは、過去の対戦記録からの機械学習だけではなく、プログラム同士の自己対戦からも学習することによって人間のチャンピオンに勝てるようになったという実績がある（Silver *et al.* [2016]）。この人工市場シミュレーション研究は、複数の知能の相互作用により生じる複雑な市場現象を機械で再現する試みであり、社会的状況における人工知能プログラムの振舞いを分析する試みでもある。同シミュレーションの成果の1つとして、東京証券取引所との共同研究による金融市場の制度検証があり、ティック・サイズ（注文価格の最小単位）変更の際の方針決定に貢献した（水田ほか [2013]）。

4. 今後の方向性と期待

本稿では、人工知能技術を用いたビッグデータ解析がファイナンス分野へと積極的に応用されている現状について、実際に提供されているサービスや研究事例を基に紹介した。

このまま金融市場でのデータ活用が高度化していき、「データに語らせる」という立場でデータを適切に分析すれば、資産運用や金融実務に必要な事実がおのずと明らかになるのであろうか。そして、金融市場の分析に際しては、人間の思考や洞察力に頼らなくても済むようになるのであろうか。そのようなことは、少なくともここ数十年では起こりえないと考える。機械によるデータ解析には、現在どうしても人間の能力にはかなわない点がいくつかある。そのため、しばらくは人間に取って代わるものではなく、人間の能力を増大させる道具として活用されることになるであろう。

現状では、全ての状況で人間に勝てる万能なデータ解析プログラムを構築することは難しい。機械学習を始めとする人工知能技術は、状況変化が少ない目先の予測は得意であり、スピードの面では人間は太刀打ちできない。しかし、経済構造の変化を含む長期的な市場分析や、政治状況や世界情勢の変化に起因する、今までに無いような新しいマクロ的な環境での市場予測は、現在の人工知能技術には困難である。囲碁や将棋のように、未来永劫ルールが変わらない世界では、人工知能が人間よりも優位であるかもしれないが、金融市場のように、その時々さまざまな社会的要因に応じてルールが変化していく世界では、まだまだ人間にしかできない作業は残されている。

今後の金融データ解析の発展の可能性は、まず、複数のデータ解析プログラムを適切に組み合わせる手法が考えられる。例えば、ティック・データ等を用いた短期的な板の動向分析と、経済ニュースのテキスト・データ分析によるイベントの影響分析、マクロ経済指標を用いた長期的な経済動向分析等のように、予測期間の長さや予測対象（相場そのもの、相場を取り巻く経済環境等）の違いによって、さまざまなモデルをデータ解析によって構築し、そのうえでそれらのモデル同士の関係性をデータに基づいて学習していくような手法が有望である。さらに、現在機械学習によるデータ解析が苦手としている金融分野の常識の獲得を、大規模なデータの解析と人間の判断モデルの統合によって改良していくことも考えられる。そして、複数の知能の相互作用により、金融市場全体の挙動を再現するような試みも、その成果が期待される。

これからの資産運用業務は、機械ができる範囲の作業（定型的な分析）は機械に任せて、機械ができない範囲の作業（長期予測、転換点での予測）を人間が自分の能力を活かしてじっくりと分析することになるであろう。人工知能技術ツールとして使いこなして、人間にしかできない課題に対して自分の能力を拡張して取り組んでいくことが求められていく。

このような新たな金融データ解析を含むフィンテック分野では、日本の現状

は残念ながら世界的にみて少し立ち後れている状況である。イギリス大蔵省が、世界の複数の国々のフィンテックの現状と将来について政策面・人材育成面・技術面等から国際比較をしたレポートでは、日本は比較対象にも入っていない (Ernst & Young [2016])。一方で、例えば英国のシティでは、世界のフィンテックの中心地になることを目標に、官民を挙げてさまざまな取組みをしている。制度上の支援や環境整備のほかに、金融機関や大学等が協力してフィンテック人材を育成するために、フィンテック・ハッカソン（フィンテック関連のプログラム開発イベント）等のさまざまなイベントを開催している。シンガポールでも金融管理局が中心となって、さまざまな金融機関やIT企業、シンガポール経営大学、シンガポール国立大学等のアカデミアを含む国中の機関を協力させ、金融サービスにおけるテクノロジー活用の促進を進めている。

日本のフィンテック分野での潜在能力は非常に高い。必要なのは、金融 (finance) と情報技術 (technology) をつなげる場と人材である。残念ながら、いわゆる文系と理系に分かれた長年の教育により、日本では金融分野と情報技術分野の間にある壁は高くなっている。お互いの分野で活躍する人材の間で、考え方や習慣、価値観、文化が異なっている。そのような壁を取り払っていくためにまず必要なことは、大がかりなことでもなくてもよいので、両分野の連携によって実務的な成功例を増やしていくことである。そうすれば、一緒に仕事をすることがお互いに必要となり、連携する機会がますます増えていくであろう。また、情報通信分野と金融分野を繋げる人材を育成する場や、さまざまな立場の人たちがオープンに協力する機会も構築されていく。本稿で紹介したさまざまな事例がそのきっかけとなる可能性を含んでおり、これからの発展を大いに期待したい。

参考文献

- 和泉 潔、「金融市場：人工市場の観点から」、杉原正顯編『岩波講座計算科学 < 6 > 計算と社会』、岩波書店、2012 年
- ・斎藤正也・山田健太、『マルチエージェントのためのデータ解析』、コロナ社、2017 年
- ・坂地 泰紀・伊藤 友貴・伊藤 諒、「金融テキストマイニングの最新技術動向」、『証券アナリストジャーナル』、日本証券アナリスト協会、第 55 巻 10 号、2017 年、28~36 頁
- 宇宙航空研究開発機構、「地球の周りを回っている人工衛星の数はいくつくらいありますか?」、ファン!ファン! JAXA!、宇宙航空研究開発機構
(<http://fanfun.jaxa.jp/faq/detail/57.html>、2017 年 3 月 6 日)
- 日本経済新聞、「AP 通信のツイッター乗っ取り 偽情報で株乱高下」、2013 年 4 月 24 日 (https://www.nikkei.com/article/DGXNASGM2402U_U3A420C1EB1000/、2013 年 4 月 24 日)
- 水田孝信・早川 聡・和泉 潔・吉村 忍、「人工市場シミュレーションを用いた取引市場間におけるティックサイズと取引量の関係性分析」、JPX ワーキング・ペーパー、Vol. 2、日本取引所グループ、2013 年
- Bollen, Johan, Huina Mao, and Xiaojun Zeng, “Twitter Mood Predicts the Stock Market,” *Journal of Computational Science*, **2**(1), 2011, pp.1–8.
- Ernst & Young, “UK FinTech: On the Cutting Edge: An Evaluation of the International FinTech Sector,” Ernst & Young, 2016.
- Fujimaki, Ryohei, Yasuhiro Sogawa, and Satoshi Morinaga, “Online heterogeneous mixture modeling with marginal and copula selection,” *Proceedings of the 17th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2011
- Liu, Enda, Tomoki Ito, Kiyoshi Izumi, Kota Tsubouchi, and Tatsuo Yamashita, “Extraction of Bi-graph Structures Among Multilingual Financial Words Using Text-Mining Methods,” in Aruka, Yuji and Alan Kirman eds. *Economic Foundations for Social Complexity Science: Theory, Sentiments, and Empirical Laws*, Springer, 2017, pp.179–191.
- Nassirtoussi, A. Khadjeh, Saeed Aghabozorgi, Teh Ying Wah, and David Chek Ling Ngo, “Text Mining for Market Prediction: A Systematic Review,” *Expert Systems with Applications*, **41**(16), 2014, pp.7653–7670.
- Raghu, Maithra, Justin Gilmer, Jason Yosinski, and Jascha Sohl-Dickstein, “SVCCA: Singular Vector Canonical Correlation Analysis for Deep Learning Dynamics and

Interpretability,” *The Thirty-First Annual Conference on Neural Information Processing Systems (NIPS 2017)*, 2017, pp.6078–6087.

Silver, David, Aja Huang, Christopher Maddison, Arthur Guez, Laurent Sifre, George van den Driessche, Julian Schrittwieser, Ioannis Antonoglou, Veda Panneershelvam, Marc Lanctot, Sander Dieleman, Dominik Grewe, John Nham, Nal Kalchbrenner, Ilya Sutskever, Timothy Lillicrap, Madeleine Leach, Koray Kavukcuoglu, Thore Graepel, and Demis Hassabis, “Mastering the Game of Go with Deep Neural Networks and Tree Search,” *Nature*, **529**(7587), 2016, pp.484–489.