

解題

証券アナリストジャーナル編集委員会

第二小委員会委員 大庭昭彦 CMA

金融のデジタル化が進み、非金融の世界に拡大していく中で、人工知能（AI、機械学習）がこれまでになかった新しい方法でビッグデータに活用されている。この新しい方法は、情報処理、確率・統計、経済学、社会科学、心理学といった多くの分野の学際的な位置にある方法であり、データサイエンスと呼ばれている。データサイエンスを行う前提として、対象がデータ化（デジタル化）され、利用のための法制度が整備されている必要がある。また、従来の方法では難しかったが扱えれば価値が大きいテーマを優先して、データサイエンスを活用していくべきだろう。今月号の証券アナリストジャーナルの特集テーマ「金融のデジタル化とデータサイエンス」では、政策当局や、大学、民間と様々な立場から、前述の目的に合った、先端的で重要な報告を集めることができた。

中島論文「金融デジタライゼーションの進展に伴う制度整備」では、主に三つのテーマでの制度整備を解説している。一つ目の金融業界のオープン・イノベーションを目指した制度整備では、銀行業界がオープンAPI（Application Programming Interface）の導入を表明すれば、電子決済等代行業者は銀行との契約と利用者同意の下で、利用者口座のID・パスワード情報なしにサービス

を行えるようになる。二つ目の金融機関の持つデータの利用促進を目指した制度整備では、金融機関業務に保有情報の第三者提供義務を追加した。これにより金融業自体へのデータ活用の促進も期待されている。三つ目の金融サービス仲介法制的整備では、銀行・証券・保険全ての金融分野のサービスを横断的にワンストップで提供する仲介業者の登録を効率化する。また、今後の課題として、顧客データ、特に個人の取引履歴データなどを活用することでプロ投資家として扱うことのできる個人の基準を精緻化する、また高齢者の中でも能力や状況に応じた対応ができるようにする、といった案が提示される。これまでのプロ投資家の基準は金融資産額や取引経験年数、高齢者対応では75歳、80歳といった年齢だけが目安になってきたことを考えれば、大きな進歩になる可能性がある。人生100年ともいわれる時代に、高齢者のお金にかかわる幸福度を高めることを考える「金融ジェロントロジー」の観点からも、後者のきめ細かな対応を促す制度整備はとても重要な課題である。

桑津論文「データ主導による非金融との境界領域をめぐる動向」では、これまで金融分野は情報との親和性を背景にデータ活用の先端分野を走っ

てきたが、非金融分野でも「実装」を目指す形で重要な深化が進んでいることを報告している。例えば、流通業では、従来のPOSデータ解析は、顧客の購買と基本属性や天候などによる売れ筋解析で、商品を主軸としていた。これに対して、ネット通販が本格普及を開始した2010年以降、顧客単位の購買履歴、販売サイト上でのページ選択などの挙動、検索サイトなどとの連動による顧客の導入ルートなど、多様な新しいデータの利用で、主軸が顧客に移動した。流通業で商品主導の解析から顧客主導の解析になってきたということがデータ主導の始まりだったということだ。製造業では特に自動車産業で、自動車保険・中古車販売のデータ主導、在庫などに対する個別の与信などが始まろうとしている。運輸業では特に陸運での応用、都市管理におけるスマートシティ、スーパーシティといった従来はSFにすぎなかった世界の現実化が、データ主導で管理せざるを得ないほどのサイズの新興の都市の出現に後押しされる形で、進むことが予測されている。最後に、「データは誰のものか？」という古くて新しい問いかけや、新しいテクノロジーがリアルな社会で「実装」されていくことでデータ主導（データサイエンス発展）が成果を上げているようだという主張は、大変示唆に富むものである。

宮川論文「ビッグデータ分析の金融実務への実装—予測と因果推論を用いた実務課題の解決—」は骨太だ。まず、ビッグデータを用いた分析の類型として「予測」と「因果推論」について詳しく解説した後、三つの分析例（事例）を挙げている。事例はそれぞれがかなり重い分析である。一つ目の事例は、企業レベルのビッグデータを用いた企業の退出・成長の予測モデルの構築である。結果として、アウトカムの種別によらず、機械学

習ベースの予測モデルの精度がロジットベースの予測モデルを凌駕していることが分かった。また、相対的に休廃業に比して倒産に関する予測の難易度が高かった。二つ目の事例は、既存の企業レベルの不正会計検知モデルを踏まえて、企業ごとの勘定科目の異常検知を試みたものである。具体的には、各企業の勘定科目レベルデータを複数年にわたって収集したデータに対してスパースモデリング手法を用いて各勘定科目間の関係性を構造学習し、その結果を踏まえて、「特定の企業×年×勘定科目」レベルでの異常度を算定し、高精度のモデルを得た。三つ目の事例として融資契約やリース契約に現れるコスト（金利、手数料）のプライシングをテーマとした因果推論手法の応用を挙げる。開発したプライシングモデルを利用して利潤を最大化することで、「最適価格」が決まってくる。この取組みのポイントは、こうした最適価格設定が、契約主体や外部環境に関する設定ごとに行い得るという点にある。これは、古くからある差別価格の議論が学生割引、シニア割引、期間キャンペーンといった粗い応用しかされてこなかったのに対して、幅広い商品群に対して精緻に適用可能なアイデアであるといえる。論文の最後で示される宮川氏の展望での、人とAIのコラボレーションの夢も面白い。例えば、予測をAIだけ、人間だけにさせるのではなく、得意なタスクをそれぞれに割り振って、結果を組み合わせるのが良いというアイデアである。この試みはまず医学分野で始まっているというが、金融やその他の領域でも似た問題、例えば、個人向けの投資アドバイスなどから進んでいくことが期待される。実際に、先行してAIだけによる投資アドバイスの提供サービスが進んでいた米国で、特に富裕層ではAIのアドバイスに人のアドバイスを組み合わせた方が顧客満足度が高いことが分かってきたことと考

え合わせると興味深い。

「METI×NOMURAコンシューマーセンチメント・インデックス・シリーズ」は、本特集の他の3論文でもビッグデータの典型として言及されているPOSデータなどを用いて消費者のセンチメント指標（インデックス）群として開発され、19年末より公開されている。中でも、同じ品目でプレミアムがついている商品とついていない商品の相対的な売れ行きを比較したプレミアム志向インデックスはユニークである。伊藤・田代・饗場論文「プレミアム品購買データから読み解く消費者心理—民間データとデータサイエンスが拓く、新しい投資・経済の視点—」では、この指標群の複雑な作成過程の紹介、指数の成分への分解、消費増税イベントでの反応、景気動向・マクロ経済指標・株価指数との関係など、幅広く分析を行っている。プレミアム志向インデックスの構成要素をプレミアムのあるブランド品の売り上げとノーブランド品の売り上げに分解すると、ノーブランド品の売り上げは冬から夏にかけて徐々に上昇し、お盆をピークに減少に転じている。これは、飲料全体の季節効果と一致する。一方で、ブランド品の売り上げには飲料平均のようなトレンドはなく、代わりにゴールデンウィーク、盆、クリスマス・正月に短期的な上昇が見られる。これは消費者が盆やクリスマスにプレミアムを好む行動を抽出反映していると考えられる。税金に対する反応の詳細も興味深い。14年、19年の消費増税時、消費者態度指数でみたセンチメントは増税時点まで下落しその後上昇した。プレミアム志向インデックスは増税までの低下は消費者態度指数と同時、あるいは先行していて、その後の回復は急であった。外的なネガティブショックに対する心理的な回復力が平均とプレミアム志向で違うと広く

解釈してみると面白い。次の章では、景気やマクロ指標などを使った分析をしている。プレミアム志向インデックスは、建設・資材、機械、自動車・輸送機、金融など、いわゆるシクリカル銘柄の多い業種と相関が強い。逆に、電力・ガス、医薬品などディフェンシブ銘柄との相関は比較的弱い。こうした株価との関係は消費者態度指数には見られなかったということで、消費者心理と投資家心理の間の新しい関係を見いだせた可能性もあり、興味深い。なお、本論文はビッグデータを用いて個人の心理を測るための前処理（指数の開発）を行い、イベントに対する反応や他の経済指標との関係を分析するという作りになっている。ビッグデータの構造が複雑なので、説得力を持った前処理自体の説明が複雑になるためだが、今後ビッグデータの活用と個人の行動予測が盛んになると、こうした構造の論文自体が増えてくるのかもしれない。

さて、四つの特集論文に共通しているキーワードとして、「個」があるといえる。個人のデータ、個別企業や個別イベントの分析について、利用可能なデータの量が増え、質が向上し、分析結果を利用者に届ける方法も個別化している。分析の質の向上とコストの低下、利用者からの価値の向上が同時に起こっていることになる。これは初めてのことでないだろうか。「個」に関連して繰り返し現れるアイデアが、広義でのパーソナライゼーション（個人化）である。狭い意味でのパーソナライゼーションといえば、インターネットにおいて個人の属性、趣味嗜好などに基づいて表示するウェブページを変化させる程度のことを指す。「今週末あなたにおすすめの映画はコレ」というやつである。ウェブページから離れて、「個人個人で異なるソリューションを提供するためのフレ

ームワーク」まで定義を広げれば、行政サービスの制度設計や民間企業の最適なマーケティング・商品のプライシングもパーソナライゼーションに含まれる。論文中でもパーソナライゼーションの事例が多く紹介されているが、今後急速な高度化と利用の拡大が期待できる分野である。また、パーソナライゼーションができるなら、全体と個人の間間的なレベル（集団）での分析や応用も可能になってくる。中間層であるセグメントの利用に関連して対象のネットワーク構造を使った分析が増えていくことも考えられる。機械学習で推定し

たモデルというとブラックボックスになりやすいのだが、ネットワーク構造を見ることは可視化して理解しやすくするための手段としても有望だ。

金融のデジタル化と利用のための法制度整備が進み（中島論文）、流通を始めとした非金融の世界に拡大していく中で（桑津論文）、データサイエンスがビッグデータに応用されている（宮川論文、伊藤・田代・饗場論文）。最近になって多くの重要な進展があったことも確認でき、今後は更に急速な利用拡大を期待したい。