

機械学習の経済分析への利用—近年の発展

大橋 和彦

(証券アナリストジャーナル編集委員会委員)

1. はじめに

現代の企業にとってデータの分析とそのビジネスへの活用は必要不可欠であり、そのための人材確保も極めて重要な課題となっている。そんな中、とある日本の有名企業（グループの一家）が機械学習・データサイエンス等の研究開発を目的に出した修士・博士向けの求人広告にあった一節、「前提スキル:数学、統計学、物理学、化学、生物学、情報科学、工学をはじめとした、数理的分野における修士以上の学位、もしくは同等以上の知識と経験」にかみつくTwitterに出くわした。曰く「経済学とかマーケティングサイエンスは明示されないですかね」と。

「機械学習、データサイエンスといった数理的分野の求人なのだからこの内容は当たり前でしょ」と思われる人も多いかもしれないが、実はこのTwitter氏のつぶやきには一理ある。というのも、ここ数年の間で、経済学Ph.D.（博士号）取得者を雇い入れる米国のテック企業（Tech Companies）が急増しているからだ。最近書かれた、Athey and Luca [2018] によるハーバード・ビジネススクールのワーキングペーパー“Economists (and Economics) in Tech Companies”には、この状況が詳しく説明されている。

彼らによれば、Amazon、eBay、Microsoft、Facebook、Airbnb、Uber等は経済学Ph.D.取得者からなるチームを抱えている。その中でもAmazonは過去5年間で150人以上の経済学Ph.D.を雇い、通常の大学の経済学部よりもはるかに多い経済学Ph.D.を抱え、しかもその規模は拡大し続けている。更に、Coursera、Expedia、Google、Netflix、Pandora、Ripple、Yelp、Zillowといった企業が経済学Ph.D.を雇い入れているとのことで、Athey and Luca [2018] にはそういった対応をする企業名が（彼らが知る範囲だけでも）50社以上挙げられている。

このようなことに彼らが詳しいのは、彼ら自身がテック企業と共同で仕事をしているもしくはした経験があり（AtheyはMicrosoft、Expedia、Ripple等、LucaはYelpやその他のテック企業）、また彼らの学生にも経済学Ph.D.取得後にこういった企業に職を求める者が増えているためである。もちろん、テック企業が彼らを雇うのは、企業にとって経済学の知見が役に立つからであり、Athey and Luca [2018] は、テック企業がどういった問題に答えを出したいと求めているか、そして経済学のスキルがそれらにどのように役立つかを説明している。

Athey and Luca [2018] が他の分野に比べ経

Amazon) が、顧客の購買行動のデータから、青汁を購入する消費者にはランニングシューズを購入する人が多いことを見いだしたとする。また、これは、健康に関心のある消費者が青汁とランニングシューズを同時に購入する傾向があったためであるとする。言うまでもなく、前者は、青汁を購入することとランニングシューズを購入することの間の(正の)相関関係であり、後者は、健康に関心があることを原因、青汁やランニングシューズを購入することを結果とする因果関係である。

この青汁とランニングシューズの相関関係を、経営者はどう利用すればよいであろうか。もしもこれを因果関係と見誤って、「青汁の売り上げを増やすことでランニングシューズの売り上げを伸ばそう。そのために青汁の価格を下げよう」というような意思決定を行ったら、おそらく失敗することになるだろう。そこに因果関係はないからだ。一方、「ランニングシューズを買いそうな消費者を見つけたい」ということであれば、青汁を購入する消費者はランニングシューズも購入する傾向があるという相関関係を利用することで、青汁の購入データからランニングシューズを買いそうな消費者を予測することはできる。即ち、相関関係は、必ずしも因果関係を意味しないが、予測には役立つのである。

こう聞いて、「いくら何でもこんな稚拙な間違いはしないでしょ」と思われる人もいるかもしれない。確かにこの例についてはその通りなのだが、現実の状況は稚拙か稚拙でないか簡単に分かるほど単純なものではない。実際には膨大なデータを分析しなければならず、そのために機械学習等の方法を用い、そこから出てくる思いがけない—しかしもしかししたら見せかけかもしれない—多

くの「関係」の中から、利用できそうなものを見いださねばならない。だからこそ、発見した「関係」が単なる相関関係なのか因果関係なのかを判別し、予測に利用するか因果関係として利用するか適切な使い分けを行えることが、データ分析の利用には重要になる。

3. 機械学習の利用法

さて、その機械学習だが、それから得られた結果を利用する際には注意を払うことが必要となる。というのも、機械学習は、大変有用な道具ではあるものの、(当たり前のことだが) 全ての問題を解いてくれるものではないからである。一言でいえば、機械学習は、データ内の変数の関係をうまく説明する関数を見いだしたり、似通った性質を持つグループにデータを分類したり、そのグループに共通する性質を見いだす技術であり、広い意味での相関関係を見つけるのは得意だが、因果関係を見つけるものではない。それゆえ、機械学習は予測には強力な道具となる一方、因果関係の理解には直接役立つものではない。

一方、伝統的な実証経済学の主要な目的は因果関係を明らかにすることであり、ここに機械学習との齟齬がある。経済分析において機械学習の手法がこれまでなかなか浸透しなかったのはこのためであるが、近年、機械学習と経済分析—とりわけ因果推論—を融合させる新たな分析方法の開発が進められている(その中心人物の一人が、先にも登場したスタンフォード大学ビジネススクール教授のSusan Atheyである)。

この点をレビューしたAthey [2018] では、経済分析への機械学習の利用を考えるに当たって、まず両者の目的の違いを議論する。機械学習は、データに含まれる諸変数の関係に適合するモ

