特　集「人工知能学会・情報処理学会共同企画―第２部「人工知能における人道とは」―」

# 機械学習の非技術的課題

## Non-Technical Issues of Machine Learning

丸山　宏　株式会社 Preferred Networks
Hiroshi Maruyama　Preferred Networks, Inc.
hm2@preferred.jp

本特集の情報処理学会側のエッセイ [丸山 16] では，近年進歩が著しい深層学習技術が，情報システム開発にどのようなインパクトを与えるかについて考察した．本学会側の本稿では，人工知能技術の社会・制度・文化的側面について考えてみよう．深層学習に関して筆者が直近の大きな問題と考える非技術的課題は，(1) 知財としての学習済みモデルと，(2) 確率的なシステムの社会受容についてである．

### 1. 知財としての学習済みモデル

人工知能に関する知財としては，政府が知的財産推進計画 2016 [内閣知的財産戦略本部 16] の中で「人工知能によって自律的に生成される創作物」の著作権について紙面をさいて議論している．もちろん，いずれは人工知能による小説や音楽などのコンテンツが現れてくるのだろうが，私達にとってもっと直近に考えなければならない知財の問題がある．それは機械学習による学習済みモデルの知財である．

訓練データから生成された学習済みモデルが，もともとの訓練データの著作権を侵害するかどうかについては，2009 年に改正された著作権法 47 条の 7 が関係するかもしれない．この条文は，そもそも，情報解析された結果は利用される著作物の表現とは異なるものであろうから，その結果をどのように利用しても自由であるとの思想に基づいており，電子計算機で「情報解析」を行う場合には著作物から二次著作物をつくってもよい，とも解釈することができる．また，この「情報解析」とは，「大量の情報から（中略）統計的な解析を行う」ことである

**著作権法 47 条の 7**
著作物は，電子計算機による情報解析（多数の著作物その他の大量の情報から，当該情報を構成する言語，音，影像その他の要素に係る情報を抽出し，比較，分類その他の統計的な解析を行うことをいう．以下この条において同じ．）を行うことを目的とする場合には，必要と認められる限度において，記録媒体への記録又は翻案（これにより創作した二次的著作物の記録を含む．）を行うことができる．

から，深層学習を用いて訓練データから学習済みモデルをつくることは，元データの著作権にかかわらず自由にやってよい，とも読める．これは機械学習の研究者にとっては福音である．

一方，学習済みモデルそのものから，さらに「統計的な解析」を加えて新しい二次モデルをつくることも可能である．この場合も，元の学習済みモデルに著作権を認めたとしても，著作権法 47 条の 7 によって，その結果をどのように利用しても自由であると解釈することができるのかもしれない．それはそれで素晴らしいことではあるが，一次モデルの知財は保護されなくてよいものだろうか，という疑問も残る．

深層学習が急速な進展を遂げているといえ，どんなデータに対しても魔法のように学習済みモデルができるわけではない．大量の訓練データを入手し，それらを選別・整理し，ニューラルネットワークの形とハイパーパラメータを決め，これまた大量の計算資源を投入して訓練を行い，さらにそれを評価する，という一連の作業には，多くの資源とノウハウが必要である．だから，二次モデルを誰でも自由につくってよいということになれば，一次モデルを作成する者にとってのインセンティブが失われ，あるいは一次モデルを厳密に秘匿することによって，学習済みモデルの再利用が妨げられるかもしれない．

特許法は本来，発明者の権利を保護することによって，発明を行うインセンティブを喚起しようというものである．しかし，現在の運用では，発明を利用する側が慎重にならざるを得ない状況も多々ある．学習済みモデルについても同様である．過剰な保護は，良い学習済みモデルの再利用を妨げかねない．学習済みモデルに限らず，良い知財は多くの人に使われてこそ人類の繁栄に資するものである．そのことを十分に考慮したうえで，議論を進めていきたいものである．

### 2. 確率的なシステムの社会受容

機械学習を取り巻く非技術的課題のもう一つとして，確率的な振舞いを行うシステムを，どのように社会が受け入れてくれるか，という問題を考えよう．

深層学習の実応用に関してよく聞く批判は，「深層学

習の結果は説明できない．このような技術は人命に関わるようなシステムには利用できない」というものである．2016年5月に起きたテスラ社の自動運転車の事故では，システムの中に深層学習が用いられていたかどうかは不明だが，画像認識など何らかの形で統計的機械学習アルゴリズムが用いられていることは想像に難くない．

統計的機械学習とは，その名のとおり，訓練データの統計的性質に基づいてシステムを構築する技術である．自動運転のように多様な状況に対処するシステムでは，起こり得る状況のすべてを数え上げられるわけではない．したがって，訓練データは，実際に起きることの確率的なサンプルでしかないことになる．確率的なサンプルであるから，偏りがあるかもしれない．もし偏ったサンプルで学習した結果を使うのであれば，出来上がった結果も偏ったものになるだろう．ポイントは，どんなに多くのサンプル点を取ったとしても，確率的なサンプリングの結果にすぎない，という点にある．であるので，システムの実行結果はやはり確率的であるとしかいえない．

確率的な振舞いを行うシステムは，社会に受け入れられないのだろうか．実は，統計的機械学習に基づかない，普通のシステム開発においても，100％の安全性を担保することはまず考えられない．一つには，ソフトウェアにはバグがつきものであるからである．どんなに慎重にテストやレビューを行ったとしても，ソフトウェアのバグを完全になくすことが不可能であるのは，ソフトウェア開発者なら誰でも知っていることである．加えて，たとえ，与えられた仕様に対して完全に仕様どおりのシステムが開発できたとしても，その仕様は常にある環境を仮定していて，その環境が変われば前提が変わるので，古い環境に適合したシステムが新しい環境で安全であるという保証はない．

このように，今までのシステム開発を行ったとしても，依然としてシステムの振舞いは多分に確率的である．違いは「確定的な結果を出すシステムをつくろうとしているのか」，「確率的な結果を出すシステムをつくろうとしているのか」という最初の意図の違いにすぎない．このような違い（または違いのなさ）を社会が受容するようにはどうしたらよいのか，頭を悩ます日々である．

## ◇ 参 考 文 献 ◇

[丸山 16] 丸山 宏：プログラミングパラダイムとしての深層学習，情報処理，Vol. 57, No. 10, pp. 974-975（2016）
[内閣知的財産戦略本部 16] 内閣知的財産戦略本部：知的財産推進計画2016，http://www.kantei.go.jp/jp/singi/titeki2/kettei/chizaikeikaku20160509.pdf（2016）

2016月7月31日　受理

## 著 者 紹 介

丸山　宏（正会員）

1983年東京工業大学大学院理工学研究科情報科学専攻修士課程修了後，日本アイ・ビー・エム株式会社入社．東京基礎研究所で自然言語処理，XML，セキュリティなどの研究．2006～09年同研究所所長．2011年より統計数理研究所教授．2016年より株式会社Preferred Networks最高戦略責任者．工学博士．