



情報学部 教授

# 加藤 毅

かとう つよし

## AI技術・産業を底上げする キーパーソン

### 学生が育つ研究室 毎年学会の賞獲得 産学連携で毒性予測プロジェクトに挑む

群馬大学情報学部の加藤毅教授研究室の快走が止まらない。

ということなのか。加藤教授は大学院時代以来、AI理論の改良に注力してきた。一人で研究を完結できてしまうものが多かった。

転機は2015年3月。自分の研究を手伝ってもらった学生に、学会で成果を発表しに行ってきたもらったら、図らずも賞に選ばれた。それがとてもうれしくて「受賞中毒」になってしまった。それ以降、なるべく学生と一緒に研究をして、学生の成果発表は、学会での受賞を最優先するという、奇妙なスタイルをとるようになった。毎年、数多くの賞を獲得。「一人の教員が運営するAI関係の研究室での受賞回数としては多い方だと思う」と語る。

学生のやる気を刺激する、潜在能力を引き出す、プレゼンテーションの方法を学ばせる——といった、優れた学生を育てる指導のノウハウを蓄積している。加藤教授こそが真のベスト・ティーチャーだ。

学生のエネルギーを糧に、自身は企業、他大学の研究者とともに大きなプロジェクトに挑んでいる。

—— どうして人工知能(AI)分野に進んだのですか。

中高生の時は、家にあったパソコンで、学校の数学で出てきたことをプログラムしたり、数字列を小さいものから順に並べる方法などの計算方法を考えたりして遊んでいました。大学に進み、コンピューターを使った計算方法がアルゴリズムという名前で体系化されていることを習い、情報科学への興味が深まりました。4年生になって文字認識を主要なテーマとしている研究室に所属しました。

—— およそ20年前ですね。当時、AIは冬の時代だったそうですね。

長年注目されながら全く成果を出せず失望され不遇の時代だったと言われています。しかし、私にとっては感動と興奮の連続。AIが実は線形代数と統計学からできていることを知り、その面白さにのめり込んでいきました。当時、日本では画像認識、音声認識、自然言語、医学の4分野にコミュニティが完全に分かれており、それぞれの分野でそのドメインに特化した方法が議論されていました。私は、今後は分野の壁が低くなり、より汎用的な方法が発達していくだろうと予想しました。“理論いじり”を中心に置きながら、いろいろな応用に視野を広げていきたいと思いました。その考え方で今日まで研究生活を送ってきました。

—— 先生の研究室の学生は元気なことで有名です。

2010年に群馬大学に着任し、当初はAI理論の改良に注力していたのですが、2015年3月に私の研究を手伝ってもらった学生に学会に成果を発表しに行ってきたら賞に選ばれました。それがとても嬉しくて、それ以降、学生となるべく一緒に研究をして、学会での受賞を意識して研究成果の発表をしてもらっています。その結果、毎年多くの賞をいただき(表1)、学生たちの推進力になっています。

## ❖ 汎用性を持たせることも重視

—— 学生を指導するうえで心掛けてきたことは何ですか。

学生と一緒に研究では何らかの応用分野に取り組むことが多かったです。主要な分野は、腎臓の顕微鏡画像解析、水中病原体予測、超音波非破壊検査、毒性予測、顔画像認



識などです。その際、重視してきたのは①その応用分野の問題を解決するために、既存のAI理論を改良すること②他の応用にも適用できるような汎用性を持たせること——の2つです。

—— 現在、先生が取り組んでいる研究の柱の一つが毒性予測プロジェクトですか。科研費基盤研究(B)や科学技術振興機構(JTS)の研究成果最適展開プログラム(A-STEP)「産学共同(本格型)」の支援を受けて、産学共同で進めていますね。

そうですね。A-STEPの課題名は「細胞科学とAI技術による毒性・機能性効果の予測システムの開発」です。新しい化学物質を開発するとき動物実験で人体への影響を調べますが、近年、動物実験に対する懸念が高まり、代替法の導入が進行中です。この代替法の一つがAIの活用です。具体的には、化学物質の分子構造を入力し毒性を予測するAI技術です。

代替法にはほかに、「in vitro技術(オルガノイド・マイクロ流体チップ)」、ヒト組織を用いるもの、侵襲度の低い動物実験があります。AI予測の強みは迅速、安価、大量スクリーニングが可能なこと。弱点は作用機序や新しい毒性メカニズムへの対応が難しいことです。

—— そこでAIを担当されているんですね。

このプロジェクトは、参画企業のUssioBIO株式会社および横浜薬科大学、早稲田大学の環境リスク管理や医療・薬理の研究者と共同で2023年10月から進めています。横浜薬科大学、早稲田大学とUssioBIOが細胞実験による

|          | 2015 | 2016 | 2017 | 2018 | 2019 | 2020 | 2021 | 2022 | 2023 | 2024 | 2025 (～4月) |
|----------|------|------|------|------|------|------|------|------|------|------|------------|
| 生物・化学    |      |      | 👑    | 👑👑   | 👑    |      |      |      | 👑    |      | 👑          |
| 水環境工学    |      | 👑👑   | 👑    | 👑    |      |      | 👑👑   | 👑👑   |      | 👑👑   |            |
| 非破壊検査    |      |      |      | 👑    |      |      |      | 👑👑   | 👑    | 👑👑   | 👑          |
| 汎用画像認識   | 👑👑   | 👑    |      |      | 👑    | 👑👑   |      |      |      |      | 👑          |
| 汎用AI要素技術 |      | 👑👑👑  | 👑    | 👑    |      | 👑    | 👑    |      |      |      | 👑          |

表1 研究室学生の年・分野別受賞歴

# Close-Up

データの取得を行い、私の研究室はデータから毒性予測を行うアルゴリズムの開発を担っています。

本プロジェクトの予算は総額1.5億円とそれなりに巨大ではありますが、それでもAI学習に十分な量のデータは収集できません。この制約を少しでも緩和するため「アンサンブル学習」と「Mixup データ拡張」と呼ばれる2つの技術を毒性予測アルゴリズムに導入することとしました。

## ——アンサンブル学習はどういうものですか。

複数のAIモデルを組み合わせて予測精度や汎用性を高めるための手法です。毒性予測プロジェクトでは、独創的な手法を開発し(メモ1)、これにより毒性予測性能の向上を確認しました。

この成果は、情報科学分野を代表するシンポジウムFIT2023で高く評価され、発表した学生がFIT奨励賞を受賞しました。

## ——Mixup データ拡張は何ですか。

AIモデルは深層モデルと浅層モデルの2種類に大別でき、前者は大量のデータがある時に利用可能です。深層モデル

の学習において、学習データを水増しするために使われているのがMixup データ拡張です。われわれは本プロジェクトにおいて、生命科学研究者でも使いやすい、完全自動学習アルゴリズム理論を開発しました(メモ2)。

この成果は、情報科学分野を代表する情報処理学会の全国大会(2025年3月)で高く評価され、発表した学生が学会の学生奨励賞を受賞しました。また、われわれの論文は2025年6月に開催されるIEEE International Symposium on Information Theory 2025にも採択されています。

## ——教員一人で運営する研究室で、毎年多くの学生が学会の賞を受賞しているところはほとんどないと思います。すばらしいです。

学生が順調に育ち、われわれの研究成果が高く評価されることは、私が研究を推進するうえでのエネルギーになっています。これからも後進の指導に励み、私自身も研究と教育のバランスを図りつつ、AI技術の向上、大学の役割の情報発信に努めたいと思います。

## MEMO 1 【毒性予測プロジェクトで開発したアンサンブル学習の概要】

- ・ ヒト多能性幹細胞を1種類の化学物質に暴露させると、あらゆる遺伝子の発現の変化を観察できる。遺伝子の個数は、タンパク質に翻訳されるものだけでも2万近く。
- ・ この研究計画の予備実験において、主成分分析と呼ばれる技術により20個の遺伝子に絞り込んで、その遺伝子からネットワークを推定し、これを、私自身がかつて開発したAIアルゴリズムの入力とすることで効果を発揮することを確認していた。
- ・ この方法だと、残りの多くの遺伝子発現データを捨てることになるが、本プロジェクトでは残りの遺伝子を用いて、複数の遺伝子ネットワークを作り、それぞれから予測した結果の多数決を取る方法を考案した(図1)。
- ・ これはアンサンブル学習の一種である。



図1 多数決で毒性予測

## MEMO 2 【毒性予測プロジェクトで開発したのはMixup データ拡張(図2)の概要】

- ・ 本研究では、ウェットの研究者でも簡単に学習できるアルゴリズムが求められていた。生物系の研究者はウェットとドライに大別されている。ウェットの研究は生物学的実験を行う研究を指し、ドライはウェット研究で得られたデータを解析する研究を指す。
- ・ これまでの深層学習モデルの学習は勾配降下法とよばれるアプローチの一択しかなかったが、この方法是对話的に学習を調整する必要があるため、ウェットの研究者では扱いづらいものだった。
- ・ 一方、浅層モデルでは、だれでも簡単に学習できる完全自動のアルゴリズムが整備されていたが、Mixup データ拡張を施した場合、完全自動のアルゴリズムの前提条件に違反する。
- ・ それは、Mixup データ拡張によって、毒性有りが無しかという離散的なカテゴリー以外に、毒性の有無の中間を表す曖昧なカテゴリーが生成されてしまう点にあった。
- ・ 加藤研究室は、こうした課題に対応し、曖昧なカテゴリーが含まれていても、高速で数値的に安定した、完全自動学習アルゴリズムの理論(図3)を開発した。



図2 Mixup データ拡張

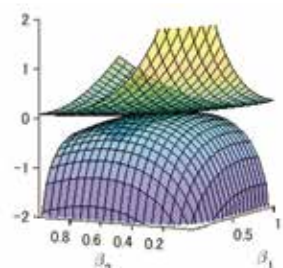


図3 完全自動アルゴリズムの理論