AI 研究センター活動報告

センター長　曹建庭

AI 研究センターは、埼玉工業大学の AI 研究拠点として人間社会との親和性の高い循環型社会に貢献できる人工知能の実現を目指しており、現実世界の様々な課題に応用できる AI 技術の開発を推進している。また、AI 関連の人材育成を行いながら産業界や地域自治体と積極に人材交流に関する活動を行う。今年度は AI 分野で日本最先端の研究機構である理化学研究所革新知能統合研究センターの研究者を招き、近年の深層学習理論とその問題点やこれからの機械学習の信頼性などについて講演会を開催した。次に AI の一つ重要な応用問題である自動診療については、東京農工大学大学院工学府の研究者を招き、癲癇の自動診断における機械学習法について講演会を開催した。

埼玉工業大学先端科学研究所

第 1 回 AI 研究センターの講演会

日　時　：　令和 3 年 6 月 24 日（木）17 時 00 分～

場　所　：　埼玉工業大学先端科学研究所 3 階会議室（発信元）

講師：　Qibin ZHAO (趙 啓斌)博士

理化学研究所　革新知能統合研究センター（AIP）

チームリーダー

https://www.riken.jp/research/labs/aip/generic_tech/tensor_learn/index.html

演題：　Trustworthy Machine Learning　（高信頼性の機械学習）

概要：　Deep learning has been developed to be the cornerstone of modern AI technology in the past decade. Since deep neural networks (DNNs) model has powerful expressiveness and capacity, the high-level representations and prediction function can be learned with high performance given sufficient high-quality data. However, DNNs are considered as a black-box system, which can only provide predictions but lack of human-centered explanations. In addition, DNNs are vulnerable to adversarial attacks with only tiny perturbations, which will lead to serious consequences in applications where reliability and safety are extremely essential such as autonomous vehicles or medical diagnosis. These issues destroy human trust on AI system.　In this talk, I will present an overview of trustworthy machine learning and especially focus on two topics including interpretable machine learning and adversarial machine learning.

【訳文】深層学習は過去 10 年間で現代 AI 技術の礎石となった技術として発展されてきた。深層ニューラルネットワーク（DNN）モデルは強い表現力と強大な可能性を持っているため、充分な高品質データを与えれば、高レベルの表現と予測機能はデータ学習により、高いパフォーマンスの結果が得られる。しかし、DNN はブラックボックス的なシステムであり、予測を行うだけで人間中心の説明ができないと考えられている。さらに、DNN はわ

ずかな干渉だけでの対抗攻撃に弱くて、自動運転や自動診断などの信頼性且つ安全性が極めて重要なアプリケーションに対して致命的な結果がもたらす。これらの問題は、AIシステムに対する人間の信頼を損なう。 本講演では、信頼性の高い機械学習の概要を説明し、特に解釈可能な機械学習と敵対機械学習の 2 つのトピックに焦点を当てる。

第 2 回 AI 研究センターの講演会

日　時　：　令和 3 年 7 月 8 日（木）17 時 00 分～
場　所　：　埼玉工業大学先端科学研究所 3 階会議室（発信元）

講師：　趙　旭陽（チョウ　キョクヨウ）博士
東京農工大学･大学院工学府　特任助教

演題：　癲癇の自動診断における機械学習法に関する研究

概要：　癲癇は脳内の神経細胞の過剰な電気的興奮に伴って発作を起こす慢性的な脳の病気である。現在、医者は長時間記録された脳波（EEG）データに基き、手動的に目視判断による癲癇を診断しているので、非常に時間がかかることと経験に依存している。医者の作業負担を軽減するために、高精度癲癇の自動診断システム開発が必要である。近年機械学習法は医療分野の診断に適用される例が多くなってきている。機械学習の発展に伴って癲癇の自動診断方法も大きく進化した。本講演では、癲癇の自動診断におけるさまざまな機械学習法を解説する。

上記の講演は、本大学の研究者や大学院生だけでなく、地域の先端科学研究所協力会の皆様も参加し、AI 技術の開発の推進、人材育成、社会に貢献などの目的を達成しました。

AI 研究センターは、埼玉県 AI・IoT・コンソーシアムに入会し、目的として県内 AI・IoT 関連分野の動向把握、情報収集、関連団体との交流・連携、埼玉県 AI・IoT プラットフォームの利用としています。今年度では、埼玉県の IoT・AI コンソーシアムへ加入している埼玉県立日高高校へ AI 教育を行い、オンラインで 2 回プロジェクトへのアドバイスなどを行いました。