

「KDD CUP 2020」入賞

2020年8月27日に、クロステック開発部の落合 桂一、田中 宏昌、サービスイノベーション部の勝丸 徳浩、川嶋 克明、川名 昭博、小泉 大輔^{†1}、酒井 俊樹、千田 拓矢^{†2}、出水 宰、福島 悠介、横山 啓太、若元 亮樹、橋本 雅人、ドコモ・テクノロジー社の鈴木 浩之の14名が、データ分析競技会であるKDD CUP 2020にて3部門において、それぞれ世界第3、4、7位に入賞し表彰されました。

KDD CUPはアメリカ計算機学会（ACM：Association for Computing Machinery）が主催するデータマイニング関連の国際会議KDD（Knowledge Discovery and Data Mining）で開かれるデータ分析競技会で、1997年、まだビッグデータやデータサイエンティストという言葉が無い時代から続く世界最高峰かつ最も歴史のあるデータ分析競技会です。

本大会は公開されたビッグデータを活用して、データ分析の精度などを競う世界大会です。2020年は6つの部門に分かれて実施され、世界中から企業、団体、個人、延べ4,000チーム以上が参加し、下記のとおり受賞いたしました。

- ・3位入賞強化学習部門：オンデマンド交通におけるタクシーの再配置AIの開発（出水、勝丸、鈴木）

本部門は、「強化学習」を活用し空車の車両をAIが再配置することで、各ドライバーの収益効率性の最大化を競うものであり、入賞した出水らのチームは先行研究を基に強化学習により時間別、エリア別の収益傾向を自ら学習し、収益を最大化させる配車パターンを提案するAIを開発しました。また、与えられた過去の配車実績や走行記録といったデータから車両運行のシミュレータを構築しAIを効率良く学習させることで、高精度かつ安定的な配車制御を行うことに成功しました。

- ・4位入賞機械学習部門：偽のデータによる攻撃と防御AIの開発（落合、田中、川名、千田、横山、小泉）

本部門は、論文の引用関係のような対象物と対象物との関係性の分類を行うAIに対し、偽のデータを紛れ込ませることでAIに誤った分類をさせる攻撃方法と、偽のデータによる攻撃を受けても正しく対象物を分類する防御方法の、2つのAI手法の開発で競うものです。

落合らのチームは、画像認識で主に用いられる、データをわずかに変化させることで誤認識させる手法を応用した攻撃手法と、対象物との関係性ネットワークのデータ用に開発されたニューラルネットワークの手法を利用した防御手法を開発しました。これらを主催者が提供する対象物数60万個のデータに適合させることで高い分類精度を維持することに成功しました。

- ・7位入賞機械学習部門：オンラインショップの商品検索で、ユーザが興味のある商品を当てるAIの開発（酒井、若元、福島、川嶋、橋本）

本部門は、オンラインショップの商品検索で、ユーザが入力した文から、ユーザがクリックする商品の画像を当てる手法の開発で競うものです。

翻訳などの言語処理で広く使われているBERT（Bidirectional Encoder Representations from Transformers）というディープラーニングの手法を改良し、画像に含まれる物体と検索文中の単語を結び付けられるAIを開発し、画像に含まれる物体の種類を表す単語の情報を用いて、検索文と画像の対応関係を学習するAIを開発しました。これらの複数のAIを使うことで、高い精度でクリックする商品の画像を特定することに成功しました。

ドコモは2016年からKDD CUPへの参加を続け、今回の入賞は2019年の優勝に続き2年連続、3回目の入賞となります。ドコモでは多数のデータサイエンティストを擁し、日頃からパートナー企業との協業の中で、AI・ビッグデータを有効活用し、さまざまな課題の解決に取り組んできたことが、今回の受賞につながりました。

本大会で評価された世界最高レベルのAI・ビッグデータ分析技術を活用し、AI・ビッグデータ活用ビジネスの拡大とともに社会課題解決の取組みを促進していきます。

†1 現在、ドコモ・テクノロジー株式会社へ出向中。

†2 現在、NTTセキュリティ・ジャパン株式会社へ出向中。

本誌に掲載されている社名、製品およびソフトウェア、サービスなどの名称は、各社の商標または登録商標。