

2023年7月18日

日立のデータサイエンティストが、世界的な AI データ分析コンペ「Kaggle」で準優勝し、Kaggle Master に昇格

株式会社 日立製作所（以下、日立）の「材料開発ソリューション^{*1}」に携わる高原渉（以下、高原）を含むチームが、世界的な AI のデータ分析コンペティションプラットフォーム「Kaggle（カグル）」で開催された「[Vesuvius Challenge - Ink Detection](#)」（以下、本コンペ）において、準優勝し金メダル^{*2}を受賞しました。高原はこれまでに Kaggle のテーブルデータ・画像データ・テキストデータのコンペにおいて銀メダルを受賞しており、それらと本コンペの結果を併せて高原は Kaggle Master^{*3}に昇格しました。

本コンペは、日本時間の 2023 年 5 月 16 日から 2023 年 6 月 15 日に開催され、1,249 チームが参加しました。与えられた課題は、火山噴火により炭化した古代の巻物の 3D 画像から、インクを検出する AI モデルを構築するというもので、この AI モデルにより、炭化の影響で解読が困難な巻物からでもテキストの復元が可能となります。AI モデル構築のための学習データが少なく、汎化性能の高い AI モデルを構築することが難しい課題でしたが、高原のチームは正則化^{*4}やデータ拡張^{*5}などを取り入れ、さらに多様な種類の AI モデルを複合することで、汎化性能を高め、高精度なインク検出を実現させました。高原の携わる材料開発の分野においても、学習データが少ない中で汎化性能の高い AI モデルの構築が求められるケースが多く、本コンペでの知見は材料開発に AI を適用していく Materials Informatics の領域で応用できるものであり、「材料開発ソリューション」の技術力強化にもつながると見込んでいます。

公共システム事業部デジタルソリューション推進部では、データ分析技術の向上のため Kaggle などのコンペティションへの参加が推進されており、高原を含むデータサイエンティストがデータ分析技術の向上に取り組んでいます。日立は今後も、データ分析への取り組みを推進し、社会・環境・経済価値の向上に貢献していきます。

*1 Materials Informatics などの技術でお客様の新材料の開発を支援するサービス。

*2 コンペの参加チーム数に応じて金メダルを獲得できるチーム数が変わり、250 チーム未満の場合は上位 10 チーム、250 チーム以上の場合は上位 10 チーム+0.2%に与えられる。今回のコンペでは、上位 12 チームに金メダルが付与された。

参考：<https://www.kaggle.com/progression>

*3 メダルの獲得枚数に応じてランクが決定し、Kaggle Master になるためには、金メダル 1 枚と銀メダル 2 枚を獲得する必要がある。2023 年 7 月現在で Kaggle の登録者数は全世界で 1,400 万人以上であり、Kaggle Master はそのうちの約 2,000 人と到達の難しい称号である。

*4 AI モデルが学習データに対して過剰に適合することを防ぐ手法。

*5 学習用の画像データを反転させるなどし、学習データ量を拡張する手法。



世界的な AI データ分析コンペ「Kaggle」で準優勝した高原渉

関連情報

[「Vesuvius Challenge - Ink Detection」に関する Web サイト](#)

[「材料開発ソリューション」に関する Web サイト](#)

以上