

Y06a GALAXY CRUISE のデータを教材とした AI 教育の試行的実践

山野井瞳 (成蹊大学), 田中賢幸 (国立天文台)

成蹊大学では文部科学省が推奨する「数理・データサイエンス・AI 教育」の一貫として、データサイエンスに関する教育を推進しており、2023 年度より「データサイエンス応用」という科目を新たに開講した。本科目は、データサイエンスの応用基礎力を習得することを目的として、機械学習および深層学習の演習を取り入れている。教材として利用するデータは、どの学部・学科の学生にもわかりやすいものであることに加え、演習には学生各自の持ち込み PC を使用するため、データ容量が大きすぎないこと、授業時間内で一通りの解析を終了できることが要件となる。開講初年度は、深層学習の演習として GALAXY CRUISE のデータを活用して、畳み込みニューラルネットワーク (CNN) による銀河の画像分類を実践した。GALAXY CRUISE は国立天文台が主催する市民天文学プロジェクトで、Hyper Suprime-Cam すばる戦略枠プログラム (HSC-SSP) のデータを用いて銀河の形態分類を行うオープンサイエンスである。深層学習のデータセットには、GALAXY CRUISE の銀河分類指標から「楕円銀河」と「渦巻銀河」をそれぞれ数百個抽出した後、HSC-SSP PDR3 より各銀河の PNG 画像を取得し、これを教師データとして準備した。深層学習フレームワークには Keras を採用した。Keras は Python で記述された深層学習のライブラリであり、比較的簡易なコードで実装できるため AI プログラミング初心者でも取り組みやすいという利点がある。畳み込み層とプーリング層を多層化した CNN モデルを適用した場合、正解率 95%以上を達成でき、単純な銀河分類であれば深層学習の教育用教材として活用できることを確認した。本講演では、実際の授業で扱った深層学習の実践手法を紹介し、その教育効果について議論する。