Summary of ACS' Machine Learning in Earth System Science workshop

The Machine Learning in Earth System Science workshop took place on the second day of the Atmospheric Science Conference. It was organised and run entirely by early career researchers and saw a high attendance. Attendees were exposed examples of current machine learning work in Earth Sciences, across all three NCAS themes. The facilitators focused on highlighting opportunities and limitations for machine learning from their own experience and published work. Tools that are available to get started and where to find them was a reoccurring topic. The workshop was structured with short talks from each of the three facilitators first, then a question and answer session, then small groups sessions, and then a finally plenary discussion.

The first of the initial short talks focused on Long-term Global Change with Dr Peer Nowack talking through his work using data-driven algorithms for predicting ozone and also work on causal inference. The second was from Dr Stuart Grange and within the NCAS Air Pollution theme, who gave an overview of meteorological normalisation using the R package rmweather which was then used as a worked example in one of the small group sessions. The final short talk was from Dr Leif Denby and was linked to the Climate and High-Impact Weather NCAS theme. Leif talked through his work using neural networks to automatically identify regimes of convective organisation in satellite images.

The facilitators' talks are hosted by RMetS on the ASC site. The worked example for rmweather from the small group session facilitated by Dr Stuart Grange is also hosted here. Some key resources mentioned during the workshop have been linked below.

For getting started with machine learning in Python, we recommend starting with the scikit-learn package

- https://scikit-learn.org/

Unlike Python's scikit-learn project, R's machine learning capabilities are not integrated in the same way so multiple packages are used to apply machine learning algorithms.

- The **caret** package, a package which acts as a consistent interface to different machine learning algorithms contained in a large number of disparate packages.
- The **ranger** package. A performant random forest package.
- The **kernlab** package. A package for applying support vector machines.
- The **neuralnet** package. A package for applying neural networks.