



Contribution ID: 156

Type: **Oral Presentation**

Machine Learning and Big Scientific Data: AI for Science

Monday, 14 October 2019 10:00 (30 minutes)

There is now broad recognition within the scientific community that the ongoing deluge of scientific data is fundamentally transforming academic research. Turing Award winner Jim Gray referred to this revolution as ‘The Fourth Paradigm: Data Intensive Scientific Discovery’. Researchers now need tools and technologies to manipulate, analyze, visualize, and manage vast amounts of research data.

This talk will review the challenges posed by the ever-increasing growth of experimental and observational data generated by large-scale facilities such as the ISIS neutron and muon facility and the Diamond Synchrotron and the CryoEM Facilities at the Rutherford Appleton Laboratory. Increasingly, scientists are beginning to use machine learning and deep neural networks both to automate parts of the data pipeline and also to find new scientific discoveries in the deluge of experimental data. In particular, ‘Deep Learning’ neural networks have already transformed several areas of computer science and research scientists are now exploring their use in analyzing their ‘Big Scientific Data’.

The talk will briefly review the advances made with deep learning neural networks and give some examples of the application of these methods to experimental data generated at the Laboratory. The talk will conclude with a discussion about possible benchmarks for scientific machine learning algorithms and their performance on HPC and GPU systems.

Primary author: Prof. HEY, Tony (Rutherford Appleton Laboratory STFC)

Presenter: Prof. HEY, Tony (Rutherford Appleton Laboratory STFC)

Session Classification: Plenary

Track Classification: Software