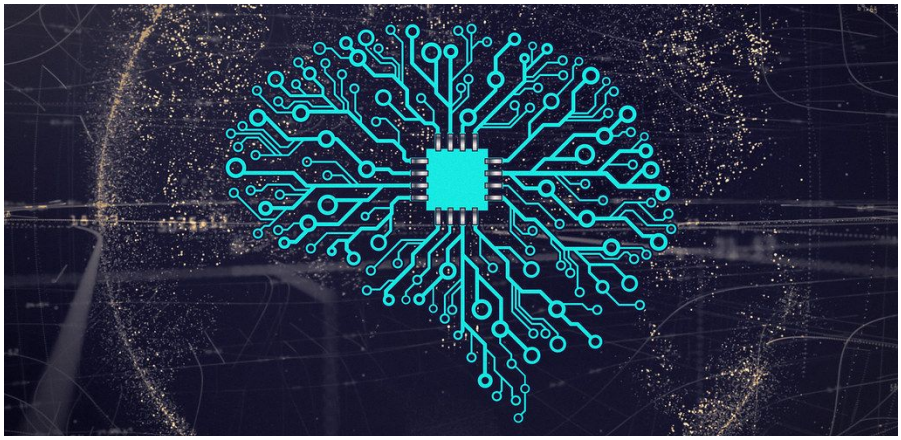
Earth Science Information Partners

Machine Learning Cluster Newsletter

March 17, 2020



Message from the Chair

Thank you for taking a look at our Machine Learning (ML) Cluster newsletter! There's a lot of ML activity in and around ESIP, listed below. Aspects of machine learning represented here include:

- the application of ML in various science domains;
- educational opportunities about a variety of ML tools;
- ML funding opportunities;
- understanding the conclusions of a trained ML model;
- agency and other organizational efforts to integrate ML: practices, strategies;

In February 2019, the American Association for the Advancement of Science (AAAS) published an article titled, "Can we trust scientific discoveries made using machine learning?" The answer was, "Not without checking." Why is that?

Rather than following explicit instructions, ML algorithms learn from data. These algorithms must always return a predication - the result of "I don't know," or "none of the above" is not an option. This can lead to uncertainty and results that may not be corroborated. Often, ML algorithms are obscure even to their creators. While this may be acceptable in some domains, such as product recommendation, the ESIP community well understands the need for transparency, understanding, and reproducibility in science.

Cultural methods to manage this concern include the development of ML standards, best practices, Technical Readiness Levels, etc., as NOAA is doing (below). NIST and even the Catholic Church are also proposing AI standards. (I have yet to find similar efforts for NASA or USGS. Anyone?)

Technical solutions can involve human-mediated construction of understandable ML models or post hoc model analysis. Here, I would particularly like to call out the upcoming **2020 Geosemantics Symposium** (below), occurring in coordination with the ESIP Summer Meeting. This one day event brings our cluster together with the Semantic Technologies Committee. This summer we focus on Explainable AI (xAI), which aims to address how decisions are made by, in our case, ML systems. Stay tuned for more details regarding date, time, and agenda.

Notable award. At the 2020 Winter Meeting in January, our Cluster Community Fellow, **Yuhan Rao,** won the ESIP Catalyst Award. That award "honors those who have brought about positive change in ESIP and inspired others to take action in the past year". I can personally attest that Yuhan, the author and instigator of this newsletter, has brought tons of positive energy (to mix physical phenomenon) to our cluster and well beyond. He has regularly inspired me to get up and do something (hopefully for the better.) Congratulations, Yuhan! And, thank you!

Anne Wilson, Ronin Institute

March 9, 2020

Inside the cluster

Highlights of ESIP-ML and cluster members' activities

- **2020 ESIP Winter meeting.** Last month's ESIP Winter Meeting has again highlighted the rapid growing interests of machine learning and artificial intelligence from the Earth sciences information community. There are multiple sessions featuring different aspects of ML/AI. For example, "AI for augmenting geospatial information discovery" (watch recording here) featured different use cases of AI technologies for agriculture, climate, and social good etc; Agriculture & Climate cluster initiated a new effort on "automated agriculture using AI" (watch recording here). Our cluster also hosted a two-session workshop focusing on the topic of training data. This workshop introduced three existing tools/platforms for generating training data -- Image Labeler, LabelImg, Bokeh, and hands-on activities to use LabelImg and Bokeh to generate labelled data (watch recording here). These three tools provide great potential to generate different types of training data via well developed open

source ecosystems. Although the winter meeting was just concluded, we have already been thinking about the topics for the summer meeting in Burlington, VT. We would love to hear your ideas and suggestions and hope to make them happen for the summer meeting.

- **2020 GeoSemantics Symposium/workshop**. Together with the Semantic Technologies Committee, we are proposing to jointly organize the summer GeoSemantic Symposium to coincide with the ESIP summer meeting. The proposed theme for this year's symposium is leveraging explainable AI (xAI) to support decision making with Earth science data. The proposed theme, echoing ESIP's strategic goal of putting data to work, aims to bring diverse stakeholders (technologists, data providers, and decision makers) together and use semantic technologies to increase our confidence in ML-supported decision making processes. At this early planning stage, we are **looking for use cases from the ESIP community** that fits the theme and develop an engaging program to put data to work!
- **AI for Agriculture (Agro AI)**. The Agriculture and Climate cluster (ACC) just started the discussion on AI for Agriculture (Agro AI) effort . We believe this can be a great opportunity to foster stronger connections and collaboration across ESIP. If you have ideas or successful examples in this area, we encourage you to share it with members from both clusters and participate in the ACC's discussion on this new effort. You can contact ***William (Bill) Teng*** (william.l.teng@nasa.gov) and ***Brian Wee*** (bwee@massiveconnections.com), ACC's co-chairs, to learn more about this exciting new effort!

All about Machine Learning

Information relevant to ESIP-ML community

- **NOAA's AI Strategy.** NOAA Research Council (NRC) has recently released NOAA's new strategies in four key science and technology focus areas. Among these key focus areas, NOAA's AI strategy aims to "dramatically expand the application of artificial intelligence (AI) in every NOAA mission area by improving the efficiency, effectiveness, and coordination of AI development and usage across the agency". NOAA's new strategy together with other federal agencies' efforts can boost Earth sciences community's effort to embrace ML/AI as tools to advance scientific discovery and support societal relevant decision making process.
- **Radiant Earth/NASA ML workshop.** During January 21-23, 2020, Radiant Earth Foundation hosted an international workshop on leveraging ML and NASA Earth observations (EO) to address environmental challenges (ML4EO). The recordings of the workshop have been archived on Radiant Earth Foundation's YouTube channel.

- **Microsoft AI for Earth Grant.** Microsoft's AI for Earth program provides cloud computing resources and AI tools for researchers and stakeholders (NGOs, local governments, etc) to solve environmental challenges. The grant provides Azure compute credits for selected projects to solve four categories of environmental challenges -- climate, agriculture, biodiversity and water. The grant has four application deadlines each year and the upcoming application deadline is **April 6, 2020**.
- **National Geographic AI for Earth Innovation Grants.** To address the many pressing scientific questions and challenges facing our planet, we must increase global understanding of how human activity is affecting natural systems and create a community of change, driven by data and cutting-edge technology. Modern technologies, such as satellite imaging, bioacoustic monitoring, environmental DNA, and genomics, can capture data at a global scale, but also produce massive, complex data sets. Artificial intelligence (AI) and cloud computing can capitalize on the potential of such data, leading to faster and more meaningful insights and creating the opportunity for transformative solutions. Learn more here. The deadline for the RFP is **July 22, 2020**.
