GC11-solidearth-20, updated on 13 May 2024
https://doi.org/10.5194/egusphere-gc11-solidearth-20
Galileo Conference: Solid Earth and Geohazards in the Exascale Era

Operationalizing deployable hazard detection technology based on machine learning

Thomas Y. Chen[1] and Luke Houtz[2]
[1]Fu Foundation School of Engineering and Applied Science, Columbia University, New York, NY 10027, United States of America (chen.thomas@columbia.edu)
[2]Arizona State University, Tempe, AZ 85281, United States of America (lhoutz@asu.edu)

As climate change advances, machine learning approaches have been developed to detect and assess impacts to infrastructure and communities by natural hazards and extreme weather events. One especially useful source of data is Earth observation and satellite imagery. For example, deep neural networks are trained on multitemporal Earth observation data to be able to perform inference on pairs of high-resolution pre-earthquake and post-earthquake imagery and output a prediction of the damage severity for any given area. Different modalities of imagery, such as infrared imagery, can be particularly useful for detecting damage. As state-of-the-art models are trained and published in the scientific literature, it is important to consider the deployability of the algorithms. Key bottlenecks to successful deployment in the exascale era include the interpretability of neural networks, computer visualization of outputs, as well as runtime dependencies of the model and memory consumption. An additional consideration is the climate impact of the significant computing resources required by large, complex models that are supposed to aid in climate adaptation. We discuss various methods of real-world deployment, including the use of drones that analyze multitemporal change in real time. Finally, we emphasize the importance of bias mitigation in machine learning and computer vision models, examining recent cutting-edge techniques like the REVISE tool, which thoroughly probes geographical biases in big data. This is required because AI requires a large quantity of data and the predictions on unseen data it makes are contingent on the data it has already seen.