Data Science Initiative

The technical areas in data science, including machine learning (ML) and its subcomponent deep learning (DL) in artificial intelligence (AI), are rapidly evolving research in statistics, computer science, mathematics, and electrical engineering. Reflecting on the larger research environments in applied data science, it has become clear that as an applied research area and a niche department, it is critical for the EAPS to keep up with and grow with the new advancements in the core areas of data science in order to ensure a leadership position in the new fields of applications. EAPS has recently developed stimulated momentum in interdisciplinary Geodata Science collaborations across campus: by creating a Geodata Science for Professionals MS degree program, hiring EAPS faculty in domain science areas with strong data science interests, successfully recruited an applied data science faculty through CoS multidisciplinary data science faculty cluster hiring, and enhanced synergies in with Math through this process as well.

To develop EAPS Data Science program, we will implement the following educational and research strategies:

Course Redesigns. We will redesign appropriate upper-level undergraduate and graduate-level courses to incorporate modules with data science applications, with the objective of providing a broader background and expertise in data science skills. This process will benefit from a collaboration with the Department of Computer Science, building on the effort to create an online graduate certificate program in Applied Data Science.

Remote Sensing Data Scientist. A remote sensing scientists that applies data science and ML/DL approaches to answer large-scale problems relevant to EAPS serve to increase scholarly work and visibility of our data science footprint. Historically Purdue has had a strong reputation in remote sensing research in Engineering, CoS, and Agriculture. EAPS will coordinate with existing expertise on campus to further attract strong candidates in EAPS remote sensing areas whose research also exhibits depth in solving data assimilation, inverse, and/or imaging problems using data-driven models to understand the physics, chemistry, and biology of terrestrial and/or planetary systems. Specifically, we recommend the department to prioritize Earth and planetary surfaces and interiors, as well as remote sensing for terrestrial and planetary atmospheres.

Physics-based Modeler. Physics-based modeling that applies recent innovations in data science and machine learning/AI for the simulation of Earth or planetary systems or components of these systems, to address high-impact and/or fundamental problems in EAPS. Prospective research areas could include data-driven applications such as data assimilation, solutions of PDE's, inversion methods, imaging, and uncertainty quantification to improve fundamental sciences in terrestrial (environmental), atmospheric, geological and geophysical, and planetary sciences.

Research Fellowship Program. Create a postdoctoral or research faculty fellowship in Applied Data Science. The EAPS Data Science program is a trans-disciplinary program integrating data science techniques with domain science applications in EAPS. Through collaborations, publications, and enhanced diversity, a postdoctoral or research faculty fellow would further promote the EAPS data science program both on- and off-campus. The fellowship would also facilitate the freedom to explore new research areas under the umbrella of the EAPS Data Science program. This could be a collective effort led by EAPS or, possibly better yet, by the College of Science, with the participation of researchers in other disciplines including computer science, statistics, and mathematics.