Training Next Generation Geophysicists to be Data Science Literate

CS303: Inspiring Metacognitive Practices

BACKGROUND

All subsurface industry fields (geothermal, oil, gas, minerals extraction, groundwater) are actively researching and implementing machine learning techniques. The new geophysics curriculum includes CS303 prepare Geophysics majors to be adept at these techniques. Our students will represent the 1st generation of geophysicists with this training. Their mastery of these topics is essential: they may not encounter many mentors on these topics in the early stages or progression of their careers.

REDESIGNED COURSE

Most data science classes are taught by computer scientists. This class is unique and useful as it is grounded in the geophysical perspective of data science. The new geophysics curriculum includes CS303 prepare Geophysics majors to be adept at these techniques. Our students will represent the 1st generation of geophysicists with this training. Their mastery of these topics is essential: they may not encounter many mentors on these topics in the early stages or progression of their careers.

Principle of Learning: Metacognition

This class is designed to explicitly engage students in the different stages of learning, with an emphasis of giving students opportunities to reflect on their learning:

- What did I learn?
- How did I learn it?

ASSESSMENT

Students will enact their own Decision Tree algorithm: asking 1 student if they like 10 given movies (training set). They will then predict if they like an 11th movie using their Decision Tree. They will be assessed on their reflection of how this can be further developed into the Random Forest algorithm. Write a minute-paper on what you took away from today’s lesson. You will need 1 “gots” (something that was clear) and 2 “needs” (something still confusing, if there are 1), addressing how we would extend the decision tree constructed here into the random forest ensemble learning algorithm.

INTENDED OUTCOMES

After completing this class, students will have the highly desired skills of data analytics (regression, classification and clustering) for the evolving subsurface industries. Since this field is ever evolving (e.g. deep learning neural networks), our students will represent the first generation of geophysicists with this training, therefore their mastery of these topics is essential, as they may not encounter many mentors on these topics in the early stages or progression of their careers.

Whitney Trainor-Guitton

Geophysics, Spatial Statistics & Value of Information, Assistant Professor

Dr. Trainor-Guitton received her PhD in Earth Energy & Environmental Sciences from Stanford University in 2010, her research interests include stochastic techniques and geostatistical modeling of geophysical data, subsurface flow modeling, carbon sequestration monitoring, and geothermal exploration.