Responsible Decisions on a Changing Planet

Coding Assignment

The problem

In this challenge the task is to predict daily river flow from precipitation and temperature measurements one day ahead. You are given a dataset “uke_eggdal_data_challenge.pkl” with the target variable “flow” in cubic meters per second at a single location and the predictors precipitation “obs_pr” in mm and temperature “obs_tas” in Celsius at 9 surrounding locations. You have to predict “flow” for the next day (e.g. 10 October 2010) from the observations of precipitation and temperature available up to the current day (including 9 October 2010). Data is given from 1970 to 2016. Imagine a hypothetical production situation where we would want to use this model for future dates.

What we are looking for

Primarily, we want to see a thorough scientific problem solving approach. Here are some things to look out for.

- Visualize the data and modeling results in illustrative ways. How can you convey your results to a non-technical audience?
- It is up to you to design the experimental setup and evaluation. How can you make sure we can trust the model you build?
- Beyond getting good predictions the code should be clearly written as well as efficient. How easy would it be for another person to understand the code?
- Likely there will be some uncertainty in the predictions. How can you provide the most useful predictions considering this uncertainty?

What to deliver

- A Python script along with a report containing figures and explanations OR an annotated Jupyter notebook.
- A short presentation (30 Minutes including questions) describing your approach.

When to deliver

Please send us the solution at least three hours before your interview slot at stephan@climate.ai

Have fun coding :)