



University  
*of* Exeter

**Project week**

**Week 11-BEM2031**

**Term2: 2023/24**

# General Assignment Feedback

- Rationale for selecting features for model
- Sometimes confusion about meaning of overfitting – overfits to data it is trained on, and poor at generalising to unseen (test) data.
- Varying gavr (or other feature) for a single observation, getting prediction probabilities, some interpretation.
- Interpretation of plot for all values of a single feature for one observation.
- Interpretation of PDP. Setting y-limits on PDP. Using subsample of data in PDP.

# Project

Submission is a single word document

Your interpretation is more important than your coding – a really coherent story about:

- What we have learned from the report
- What gaps there are in the report's analysis
- How you have filled those gaps a little bit with your visualisation/model
- What we have learned from your visualisation/model that we didn't know before
- What else could be done to address the problem?

# Project

Be specific, not general

Don't overstate

Don't understate

You don't need to solve the problem – just provide a bit of new information

That new information may or may not be actionable but it should come from your critique (what gaps are there?) and should be critiqued (what have we learned -or not- from this?)

# Project

## Structure:

In your intro/business understanding – point to where your critique/analysis will go. If you are focusing on trying to understand workload a bit more, say that in your introduction.

You can add your visualisation in the data exploration section (or model section) and your model in the model section if that makes most sense to you.

OR you can add both at the end if it makes more sense to do it that way.

There is not right or wrong answer