

DESE50001 – Data Science
Dyson School of Design Engineering, Imperial College London

Application Portfolio 3

Assignment: Modeling analysis

Each student on your team should choose a question/problem that they can answer using one of the predictive methods introduced in class.

Keep in mind that this **project is solely focused on predictive analysis**.

Your question/problem should be **very focused and specific**, and the **context clear and well described**.

When you have picked a question/problem, consider:

- Why is this question/problem interesting and important?
- What type of question/problem is it (e.g., regression, classification)?
- Will the data I have enable answering adequately this question/problem?
- Is it really a prediction problem? In which context, could I need to make a prediction, using which data, and what will I do with my prediction?
- Can I answer this question using the tools we have seen in class?

The type of question/problem you address defines the type of predictive tool you will use. For example, if you want to predict the next price movement of a stock, a regression model may be best suited. If you want to predict whether the next unseen images is a cat vs. a dog, a classification algorithm may be best (more on this later).

Start thinking about the type of model that may be most suitable to address your question/problem of interest.

Thinking ahead about your analysis and future results, make sure you can explain and justify what you are doing (why are you using this method?), that you know how to interpret your results correctly (what is the meaning of the coefficients, confidence intervals, R^2 , etc.), and what you can do to manage model biases, outliers, and dataset imbalance (more on this from Week 6).