Data Science Toolbox Portfolio Questions

05 Supervised Learning and Ensembles

Daniel Lawson — University of Bristol

Block 4

Portfolio 05

Choose **one question** and write up to **one page** about it. You are free to conduct further experiments to add weight to your results, and any additional material you generate can be submitted as an appendix. See The Assessment Page for advice.

These questions may make reference to the content from the current block.

Question R05.1: Imagine that you are a data-science consultant who has provided the analysis in Block 5 Workshop to a company. Give a 1-page "executive summary" of your conclusions, being sure to address: a) predictive performance, b) computational concerns, and c) interpretation. You are welcome to include additional analyses in the appendix, and reference figures, as appropriate.

Question R05.2: Your boss tells you to perform Bagging as it will improve performance. Explain what bagging can do for you, as well as identifying the limitations and costs that it has.

Question R05.3: Read a paper on the Quadratic Programming problem for SVM's such as Multiplicative Updates for Nonnegative Quadratic Programming in Support Vector Machines to understand and summarise quadratic programming, and explain one method for finding a solution to it. What problems can it solve, and what can it not handle? Emphasise the challenges that the quadratic component creates over simpler problems.