Response 1
Science Theory and Research Methodology
DD2205

Jim Holmström 890503-7571
jimho@kth.se

November 9, 2012

This is a response to the chapter 'The Scientific Process' from 'Advice to a Young Scientist' by Peter Medawar. It can be summarized as going trough and discussing the aspects of a scientific process at different levels, such as; testing hypotheses by experimentation, using logic systems for deduction, and the self-correcting process of critical evaluation.

It seems by the text that machine learning in many ways and at many levels tries to mimic the scientific process by, for example: setting up and testing hypotheses by experimentation (collecting data) and then refine them according to the data (which can lead to overfitting the model, see below), and repeating this process until it reaches a good enough model of the data. One problem with this search for a solution is the very vastness of the most general hypothesis space. Where humans uses the "hunch" to guide the hypothesis search the algorithms fall short by lacking the possibility of using this concept and instead needs to use a constructed subspace.

The chapter also touches on the subject of trying to fit a hypothesis to seen data and not the other way around. Doing it the other way around can be problematic especially when your working with sparse data since it is then more likely for the hypothesis under test to be perceived as good solely by the effect of coincidence, that is your hypothesis has fitted the noise from the data.