

These lectures represents a challenging question: “How do we chose what variables to include in a regression model?”. Sadly, no single easy answer exists and the most reasonable answer would be “It depends.”. These concepts bleed into ideas of machine learning, which is largely focused on high dimensional variable selection and weighting. In the following lectures we cover some of the basics and, most importantly, the consequences of over- and under-fitting a model.