

P07 – 4th Reading Assignment [Optional]

1. The Do's and Don'ts of Data Mining (Hinman) 2014

Read the blog post “The Do's and Don'ts of Data Mining” from Heather Hinman and take notes while reading (of course also possible in form of side notes / underlining in the paper itself).

2. Machine Learning Done Wrong (Chu) 2015

Read the blog post “Machine Learning Done Wrong” by Cheng-Tao Chu and take notes while reading (of course also possible in form of side notes / underlining in the paper itself).

3. 11 Clever Methods of Overfitting and how to avoid them (Langford & Piatetsky-Shapiro) 2015

Read the blog post “11 Clever Methods of Overfitting and how to avoid them” by Langford and Piatetsky-Shapiro and take notes while reading (of course also possible in form of side notes / underlining in the paper itself).

4. Consolidation

Bring what you have read into synchronization with what you've heard in lectures V06 and V07: In which ways does this new material enlarge on / contradict / extend the best practices you already studied? What do you find helpful on a first glance? What is most astonishing to you?