**Protocol**

* One-page
* Compare at least three different prediction models
* Use several complementary performance measures
* Explore at least three dimensionality reduction mechanisms
  + Filter appraoches & model-based approaches (RF based importance value)
  + Should be compared with base line models
    - Baseline model: model without any feature selection

**Report**

* Max 4 pages/task
* Small paragraph with who wrote wat part
* Reference list

**Code**

* Comments

**Big data report**

* Max 2 pages
* Only for one task

1. Write on which elements should be changed/added when the data size increases.
2. What about visualization and automation
3. What are the gains/loses when using simple (aka less computational approaches?)
4. What methods should be used when n >>> p (compute power is abundant)
5. Present a simple parallelized version of the data: divide the data in smaller subsets and analyze these in parallel. Merge the results back together and answer based on the summarized results. Discuss the process of how to divide and recombine.