Machine Learning II: Assignments #1
14 performance points (max),
email: PDF+code to jan.nagler@gmail.com
due: Tue, April 21, 2020

1. Covid-19 modelling warmup

   The SIR model is a 3-compartment model. Extend this model to 4 compartments, where the 4th compartment is for deaths (D). Mortality is modelled by new transitions from $I \to D$ taking place with death rate μ. Susceptible and Recovered do not die.

   (a) Derive the corresponding system of equations for S, I, R and D. E.g., $\frac{\mathrm{d}R}{\mathrm{d}t} = \mu I$ but this is not the only difference to SIR.

   (b) Assume that the basic reproduction number R_0 is not exactly known but only the range $R_0 \in [1.1, 6.0]$. Assume that the mortality rate μ is also not exactly known but only the range $\mu \in [0.1\%, 4\%]$. Study numerically how these parameter uncertainties affect the longterm prediction of D? Your choice how systematic you do this.

   (c) Study numerically the effects of lockdowns (by varying β), where the focus is on $D(t)$, in particular for $t \to \infty$.

   (d, optional) Can you find a way to plot the effective reproduction number as a function of time, given otherwise fixed parameters ?

   Free choice! Find suitable initial conditions $(S(0), I(0), R(0) = D(0) = 0)$, and parameters (if independent of other assumptions) for (a-d).

2. Principal Component Disasters
   Create labeled surrogate data sets. Perform a PCA/Class prediction with ovr logistic regression analysis as developed in the lecture.

   (a) 4 blobs: Create clearly separable 4-blobs but also a 'disaster' realization with strong overlaps. Study, show and compare elbow plots and prediction boundaries.

   (b–optional) 2 touching parabola spreads as shown in the lecture, but in 3d (not 2d). Study and show elbow plot and prediction boundaries.