

T.61-6070 Special Course in Bioinformatics I  
Modelling of Proteomics Data

Exercise questions on

Seminar presented by: Gopal Peddinti

Topic: Identification of protein signals from MALDI-TOF MS spectra

Date: 10.04.2008

....

1. What are Biomarkers? Imagine that you are a biologist interested in finding out protein biomarkers relevant to a particular human disease 'X'. Let us say, you would like to set up some MALDI-TOF MS analysis based proteomics experiments. Explain various tasks involved in performing the biology experiments and computational analysis in the process of finding a set of biomarkers. You need not specify experiment protocols or computational methods you would use (unless you feel more comfortable specifying them than omitting). It is enough to give an outline of tasks along with a brief explanation as to why to do it (e.g. mention things like sample collection, peak picking and various other pre-processing steps, and steps involved in higher level analysis). To justify the topic of presentation you should in particular mention the pre-processing steps well.
2. Gather how sensitivity and specificity are defined (in the context of binary classification hypothesis tests). Assume that we performed some pre-processing and detected 48 peaks in total from all IC waveforms corresponding to the synthetic MALDI-TOF MS spectra (see the paper or presentation). Further assume that only 32 of these 48 peaks correspond to the peaks that were used to generate the synthetic data. Now compute sensitivity and specificity to assess how well the method worked.