IPM-POLYTECHNIQUE-WPI Workshop on Bioinformatics and Biomathematics,
April 11-21, 2005, IPM, Tehran

Molecular Diagnosis

(Part II)

F. Markowetz
Department of Computational Molecular Biology
Max Planck Institute for Molecular Genetics
Berlin, Germany

The second lecture on classification will deal with feature selection (filter and shrinkage methods); classi- fication by Nearest Shrunken centroids (Tibshirani et al., 2002); model selection and assessment (Hastie et al., 2001, ch.7); pitfalls in model selection (Simon et al., 2003; Ambroise and McLachlan, 2002); and interpretation of results (Ein-Dor et al., 2005).

References:

[1] **Ch. Ambroise and G. J. McLachlan.** *Selection bias in gene extraction on the basis of microarray gene-expression data.* Proc. Natl. Acad. Sci. USA, **99**(2002), 6562-6566. URL http://dx.doi. org/10.1073/pnas.102102699.

[2] **L. Ein-Dor, I. Kela, G. Getz, D. Givol, and E. Domany.** *Outcome signature genes in breast cancer: is there a unique set?* Bioinformatics, **21**(2005),171-178. URL http://dx.doi.org/10. 1093/bioinformatics/bth469.

[3] **T. Hastie, R. Tibshirani, and J. Friedman.** *The Elements of Statistical Learning.* Springer, 2001. URL http://www-stat.stanford.edu/ tibs/ElemStatLearn/.

[4] **R. Simon, M. D. Radmacher, K. Dobbin, and L. M. McShane.** *Pitfalls in the use of DNA microarray data for diagnostic and prognostic classification.* J Natl Cancer Inst, **95**(2003), 14-18.

[5] **R. Tibshirani, T. Hastie, B. Narasimhan, and G. Chu.** *Diagnosis of multiple cancer types by shrunken centroids of gene expression.* Proc Natl Acad Sci USA, **99**(2002), 6567-6572. URL http://dx.doi.org/10.1073/pnas.082099299.