

**STANFORD UNIVERSITY**  
**DEPARTMENT OF STATISTICS**  
**DEPARTMENTAL SEMINAR**

4:15 p.m., Tuesday, February 21, 2006  
Sequoia Hall Room 200  
(Cookies at 3:45 in 1st Floor Lounge)

*Junhui Wang*  
School of Statistics  
University of Minnesota

**Margin-based Semi-supervised Learning**

In classification, semi-supervised learning occurs when a large amount of unlabeled data is available with only a small number of labeled data. In such a situation, how to enhance predictability of classification through unlabeled data is the focus. In this talk, I will present a novel margin-based semi-supervised learning methodology, utilizing grouping information from unlabeled data, together with the concept of margins, in a form of regularization controlling the interplay between labeled and unlabeled data. In particular, I will discuss three aspects: (1) the idea and methodology development; (2) computational tools; (3) a statistical learning theory. Numerical examples will be provided to demonstrate the advantage of our proposed methodology, particularly against SVM with labeled data alone as well as transductive SVM.