Sample Course Outline, COMP / ELEC / STAT 502  
Artificial Neural Networks and Information Theory I.  
Approximately 12 x 3 = 36 lecture hours, 3 credits

1. Introduction  
   • What is an ANN, defining characteristics  
   • Categories of ANN paradigms  
   • Learning, adaptation, intelligence, learning rule categories: supervised / unsupervised / reinforcement  
   • Application areas, history  
   • Major ANN simulation software, major journals and literature sources  
   • Hardware ANNs

2. Review of Information Theory and Statistics  
   • Gaussian and uniform distribution, covariance, correlation, moments  
   • Conditional probability, least squares, maximum likelihood  
   • Quantification of information, entropy, joint and conditional entropy, mutual information, Kullback-Leibler divergence  
   • Principal Components, subspaces

3. Associative Memory  
   • Memory, autoassociation, heteroassociation  
   • Memory matrix, recall and crosstalk  
   • Bi-directional autoassociative memories, recall from partial and noisy samples  
   • Stability of bi-directional memory

4. Simple Supervised Learning  
   • Perceptron, linear separability, XOR problem, linear and non-linear neurons  
   • Error descent, Delta-rule  
   • Proving the convergence of learning; Lyapunov’s direct method

5. The Multilayer Perceptron (MLP)  
   • The Backpropagation algorithm (BP)  
   • MLPs are universal approximators: theorems  
   • Convergence, local minima  
   • Speeding up the learning with momentum  
   • Number of hidden units, weight pruning  
   • Training concerns: generalization vs memorizing, overtraining, number of training samples, stopping criteria; scaling of inputs and outputs, preprocessing of data  
   • Function approximation, prediction, classification with MLPs

6. Unsupervised Learning  
   • Hebbian learning, stability, weight decay,  
   • Oja’s PCA nets, Sanger’s Generalized Hebbian Algorithm, Földiák method, negative feedback  
   • Hebbian learning and Information Theory – PCA and subspace connections; regression and minor component analysis  
   • Competitive learning  
   • Self-Organizing Maps (SOMs); visualization and information extraction from SOMs  
   • Learning Vector Quantization (LVQ), Adaptive Resonance Theory (ART)  
   • SOM applications, WEBSOM, PICSOM  
   • Grossberg star, anti-Hebbian learning
7. Recurrent Nets
   • Hopfield networks
   • Boltzmann Machine
   • Simulated Annealing

8. More on Speeding Up Supervised Learning and Structure Optimization
   • Radial Basis Functions, Error Descent, QuickProp
   • Conjugate Gradients
   • Cascade Correlation

9. Objective Function Methods in ANN Learning
   • Backpropagation and PCA
   • Cross-entropy, maximum mutual information (I-Max), maximum correlation (Canonical Correlation Analysis) as objective functions; use of contextual information

10. Identifying Independent Sources (Blind Source Separation) with ANNs
    • Competitive Hebbian learning; Anti-Hebbian and competitive learning; sparse coding
    • Multiple cause models, Factor Analysis
    • Non-linear PCA as an extension to Oja’s Subspace Algorithm (Hebbian learning)

   Time permitting:
   • Predictability minimization
   • The use of noise
   • Probabilistic models

11. Time permitting: Independent Component Analysis
    • Independent Component Analysis – definition of the problem
    • Information maximization
    • Projection Pursuit – maximally ‘interesting’ projections (non-linear)

Exercises/home works, exam and course project will involve programming in the student’s choice of Matlab, C or R (Fortran is also acceptable).