Sample Course Outline, COMP / ELEC / STAT 502
Neural Machine Learning I.
(Learning with Artificial Neural Networks)
Approximately 12 x 3 = 36 lecture hours, 3 credits

1. Introduction
- What is an Artificial Neural Network (ANN), defining characteristics
- Categories of ANN paradigms
- Learning, adaptation, intelligence, learning rule categories: supervised / unsupervised / reinforcement
- Major ANN simulation software, major journals and literature sources

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
- Hebb's rule; Memory matrix, recall and crosstalk
- Bi-directional autoassociative memories, recall from partial and noisy samples
- Stability of bi-directional memory
- Learning the weights with error correction
- Lyapunov's direct method for proving convergence / stability

4. Simple Supervised Learning
- Perceptron, linear separability, XOR problem, linear and non-linear neurons
- Error descent, Delta-rule
- Proving the convergence of learning

5. The Multilayer Perceptron (MLP)
- The Backpropagation algorithm (BP)
- MLPs are universal approximators: theorems
- Convergence, local minima
- Speeding up the learning with momentum
- Structural considerations: 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. Evaluation of Learning and Generalization Performance
- Selection of training and test data sets
- K-fold cross-validation
- Accuracy assessment

7. Unsupervised Learning
- Hebbian learning, stability, weight decay
- Oja's PCA nets, Sanger's Generalized Hebbian Algotrithm
- Földiák method, negative feedback (later)
- Competitive learning

- Self-Organizing Maps (SOMs); visualization and information extraction from SOMs
- Learning Vector Quantization (LVQ)
- SOM applications, examples WEBSOM, PICSOM
- SOM as hidden layer in supervised schemes

8. Recurrent Nets

- Hopfield networks
- Boltzmann Machine (time permitting)
- Simulated Annealing
-

9. More on Supervised Learning

- Vapnik-Cervonenkis Dimension for sufficient training samples
- NN outputs vs Bayes decision

10. More on Speeding Up Supervised Learning and Structure Optimization

- Network Growing with Cascade Correlation
- Cross-entropy for classification with BP
- Conjugate Gradients
- Radial Basis Functions, QuickProp

11. Objective Function Methods in ANN Learning

- Backpropagation and PCA
- Maximum mutual information (I-Max), maximum correlation (Canonical Correlation Analysis) as objective functions; use of contextual information

12. 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

13. *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).