

Syllabus
Psychology 599 - Special Topics
Introduction to Neural Networks
Bosco S. Tjan, Ph.D.
Spring 2002

Introduction to the principles and techniques of constructing large-scale parallel distributed processing models in neural and cognitive sciences.

Topics discussed will include:

Levels of analysis, representation, linear algebra, numerical techniques, Hebbian rule, linear-associators, multi-layer nonlinear networks, back-prorogation, Hopfield networks, Boltzmann machines, self-organizing maps, radial-basis function networks, support-vector machines, principal-component analysis, and information optimization.

Application areas include:

Visual perception, language, learning, memory, and social interactions.

Readings will include selected chapters from:

1. Anderson, J.A. *An Introduction to Neural Networks*. MIT Press
2. O'Reilly R.C. & Munakata, Y. *Computational Explorations in Cognitive Neuroscience*. MIT Press
3. Arbib, M.A. *The Handbook of Brain Theory and Neural Networks*. MIT Press
4. Haykin, S.S. *Neural Networks: A comprehensive Foundation*. Prentice Hall

Exams: One midterm, no final

Final project may be either an implementation of a neural network model or a term paper on a specific topic in neural network modeling.

Grade based on: class presentation - 25%, midterm - 15%, homework and programming assignments - 30%, final project - 30%.

Class will meet once a week for 2 hours, followed by 1.5 hours of lab.

Lecture Plan:

❖ **Student presentation at the beginning of the *next* lecture**

- 1/9 Introduction, levels of cognitive modeling, artificial neurons, representation, linear algebra, Matlab
- Kohonen T, 1997, *Self-Organizing Maps*, Ch. 1, Springer.
 - McCulloch WS, & Pitts W, 1943, "A logical calculus of the ideas immanent in nervous activity", *Bulletin of Mathematical Biophysics*, 5:115-133
- 1/16 Linear algebra (cont.), linear associator, autoassociator, Hebbian learning rule
- Anderson JA, 1996, *An Introduction to Neural Networks*, Ch. 6, MIT Press
 - Pitts W, & McCulloch WS, 1947, "How we know universals: the perception of auditory and visual forms", *Bulletin of Mathematical Biophysics*, 9:127-147
- ❖ Anderson JA, Silverstein, JW, Ritz, SA, & Jones, RS, 1977, "Distinctive feature,

- categorical perception, and probability learning: some applications of a neural network”, *Psychological Review* 84:413-451.
- 1/23 Non-linear activation function, single-layer perceptron, delta rule, perceptron and Bayes classifiers
- Anderson JA, 1996, *An Introduction to Neural Networks*, Ch. 8, MIT Press.
 - Haykin S, 1999, *Neural Networks: a Comprehensive Foundation*, pp 135-155.
 - ❖ Widrow B, & Hoff ME, 1960, “Adaptive switching circuits”, *1960 IRE WESCON Convention Record*, New York: IRE, pp. 96-104.
- 1/30 Multilayer perceptron, backpropagation, general practices for network training and testing.
- Anderson JA, 1996, *An Introduction to Neural Networks*, Ch. 9, MIT Press.
 - Haykin S, 1999, *Neural Networks: a Comprehensive Foundation*, §4.1-4.9, Prentice Hall.
 - ❖ Nair J, Nair SS, Kashani JH, Reid JC, Rao VG, 2001, “A neural network approach to identifying adolescent adjustment”, *Adolescence* 2001 Spring;36(141):153-62.
 - ❖ Sejnowski TJ, & Rosenberg CR, 1986, “NETtalk: a parallel network that learns to read aloud”, *The Johns Hopkins University Electrical Engineering and Computer Science Technical Report*, JHU/EECS-86/01.
- 2/6 No class
- 2/13 Radial-basis function networks and regularization theory
- Haykin S, 1999, *Neural Networks: a Comprehensive Foundation*, Ch. 5, Prentice Hall.
 - Anderson JA, 1996, *An Introduction to Neural Networks*, Ch. 13, pp 452-459, MIT Press.
 - ❖ Poggio T, & Edelman S, 1990, “A network that learns to recognize three-dimensional objects”, *Nature*, 343:263-266.
- 2/20 Kernel methods, support vector machine
- Haykin S, 1999, *Neural Networks: a Comprehensive Foundation*, §6.1-6.4, Prentice Hall.
 - Anderson JA, 1996, *An Introduction to Neural Networks*, Ch. 13, pp 452-459, MIT Press.
 - ❖ Vapnik VN, 1995, *The Nature of Statistical Learning Theory*, §5.6.4-5.8, Springer.
- 2/27 Mixture of experts, hierarchical mixture of experts, EM algorithm [language processing]
- Haykin S, 1999, *Neural Networks: a Comprehensive Foundation*, §7.6-7.14, Prentice Hall.
 - ❖ Joanisse MR, & Seidenberg MS, 1999, “Impairments in verb morphology after brain injury: a connectionist model”, *PNAS* 96:7592-7597.
 - ❖ Pomerleau D., 1991, “Efficient Training of Artificial Neural Networks for Autonomous Navigation”, *Neural Computation*, 3(1):88-97.
- 3/6 Midterm
- 3/20 Unsupervised learning, principal components analysis
- Haykin S, 1999, *Neural Networks: a Comprehensive Foundation*, §8, Prentice Hall.
 - Anderson JA, 1996, *An Introduction to Neural Networks*, Ch. 9, pp267-271, MIT Press
 - ❖ Olshausen BA & Field DJ, 1997, “Sparse coding with an overcomplete basis set: a strategy employed by V1?” *Vision Res.* 37:3311-15.
 - ❖ Linsker R, 1986, “From basic network principles to neural architecture: emergence of spatial-opponent cells”, *PNAS*, 83(19):7508-12.
 - ❖ Linsker R, 1986, “From basic network principles to neural architecture: emergence of orientation-selective cells”, *PNAS*, 83(21):8390-4.
 - ❖ Linsker R, 1986, “From basic network principles to neural architecture: emergence of orientation columns”, *PNAS*, 83(22):8779-83.
- 3/27 Self-organization map
- Haykin S, 1999, *Neural Networks: a Comprehensive Foundation*, §9, Prentice Hall.
 - Kohonen T, 1997, *Self-Organizing Maps*, 2nd Ed., Springer

- Kohonen T, 1982, “Self-organized formation of topologically correct feature maps”, *Biological Cybernetics*, 43:59-69.
 - ❖ von der Malsburg C, 1973, “Self-organization of orientation sensitive cells in the striate cortex”, *Kybernetik*, 14:85-100.
 - ❖ Thomas E, Van Hulle MM, Vogels R, 2001, “Encoding of categories by noncategory-specific neurons in the inferior temporal cortex”, *J Cogn Neurosci*, 15;13(2):190-200.
- 4/3 Information-theoretic models, entropy, Kullback-Leibler divergence, mutual information, independent component analysis, blind source separation
- Haykin S, 1999, *Neural Networks: a Comprehensive Foundation*, §10, Prentice Hall.
 - Bell AJ, & Sejnowski TJ, 1995, “An Information-Maximization Approach to Blind Separation and Blind Deconvolution”, *Neural Computation*, 7:1129-59.
 - ❖ Bell AJ, & Sejnowski TJ, 1997, “The ‘Independent Components’ of Natural Scenes are Edge Filters”, *Vision Research*, 37:3327-38.
- 4/10 Statistical mechanics, Gibbs distribution, Boltzmann machine, mean-field approximation
- Anderson JA, 1996, *An Introduction to Neural Networks*, Ch. 12, MIT Press.
 - Haykin S, 1999, *Neural Networks: a Comprehensive Foundation*, §11, Prentice Hall.
 - ❖ Geman S, Geman D, 1984, “Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images”, *IEEE PAMI*, 6:721-741.
 - ❖ Dayan P, Hinton GE, Neal RM, & Zemel RS, 1995, “The Helmholtz machine”, *Neural Comp.* 7:889-904.
- 4/24 Neurodynamics, attractors, Hopfield net, recurrent network, [memory (revisited)]
- Anderson JA, 1996, *An Introduction to Neural Networks*, Ch. 12,15, MIT Press.
 - Haykin S, 1999, *Neural Networks: a Comprehensive Foundation*, §14, Prentice Hall.
 - ❖ Rolls ET, & Stringer SM, 2001, “A model of the interaction between mood and memory”, *Network*, 12(2):89-109.
 - ❖ Rao RP, & Ballard DH, 1997, “Dynamic model of visual recognition predicts neural response properties in the visual cortex,” *Neural Comput.*, 15;9(4):721-63.
- 4/17 Project representation / general discussion