

CN550 Spring 2009**Neural and Computational Models of Recognition, Memory, and Attention**

Instructor: Gail A. Carpenter, Department of Cognitive and Neural Systems,
Boston University, 677 Beacon Street, Room 303, Boston MA 02215.
(617) 353-9483 Office hours: By appointment. gail [at] cns.bu.edu

Teaching Fellow: Ben Chandler Room 101 bchandle [at] cns.bu.edu
Office hours: Monday noon (before class); and by appointment.

Classes: Mondays, 1:00 PM - 5:00 PM, 677 Beacon Street (Basement)
January 26 – April 27, 2009

Schedule notes: There will be no class on February 16 or April 20 (holidays) or on March 9 (spring break).
CN550 will meet Tuesday Feb 17 and Thursday April 23 (BU Monday schedule these days).

Course description: CN550 develops neural network models of how internal representations of sensory events and cognitive hypotheses are learned and remembered, and of how such representations enable recognition and recall of these events. Various neural and statistical pattern recognition models, and their historical development and applications, are analyzed. Special attention is given to stable self-organization of pattern recognition and recall by Adaptive Resonance Theory (ART) models. Mathematical techniques and definitions to support fluent access to the neural network and pattern recognition literature are developed throughout the course. Experimental data and theoretical analyses from cognitive psychology, neuropsychology, and neurophysiology of normal and abnormal individuals are also discussed. Course work emphasizes skill development, including writing, mathematics, computational analysis, teamwork, and oral communication.

Class project: CN550 includes a class project, as described in the accompanying materials. Part of each class is devoted to discussion of the class project and planning for the coming week. Each student will work in a group with approximately three other students. Groups should plan to meet during the weekly discussion session and at other times, as needed.

Computational workshops: Each class will conclude with a computational workshop.

Course materials:

Lecture notes will be given out in class and made available as PDF files on the wiki.

See p. 3 for textbook information.

Other readings are available from the CN550 wiki, on each Lecture page. Please use ACS to print articles.

Evaluations: Your grade will be based on work totaling 400 points:

60	reading journals	weekly
30	article & lecture essay	due March 30
70	class project	weekly; essay due May 11
70	computational workshops	weekly
70	midterm exam	March 23
100	final exam	May 11

Your CN550 grade will also reflect your overall performance, progress, and participation.

Syllabus: http://cns.bu.edu/cn550/550_Syllabus_2009.pdf

CN550 wiki: <http://cns.bu.edu/cn550/>

CN710 wikis: Fall 2008 <http://cns.bu.edu/cn710/Fall2008/>

Fall 2007 <http://cns.bu.edu/cn710/Fall2007/>

Fall 2006 <http://cns.bu.edu/cn710/Fall2006/pmwiki.php?n=Main.HomePage>

Spring 2006 <http://cns.bu.edu/cn710/Spring2006/pmwiki.php?n=Main.HomePage>

CN550 LECTURES and EXAMS

- | | | |
|------------------|-----|--|
| Jan 26 | 1. | Overview, history, philosophy, benchmark database studies |
| Feb 2 | 2. | Supervised learning methods: Memory-based algorithms (KNN), model-independent supervised learning methods (validation & cross-validation, c-index, ROC curves, resampling, combining classifiers, component analysis), statistical pattern recognition |
| Feb 9 | 3. | Unsupervised learning: Clustering (leader, K-means), competitive learning, ART |
| Feb 17 (Tues) | 4. | Dimensional analysis, competitive networks, phase plane analysis |
| Feb 23 | 5. | ARTMAP |
| March 2 | 6. | Associative memory networks: Back propagation, multi-layer perceptrons, radial basis functions, cascade-correlation, higher-order networks |
| March 16 | 7. | Support vector machines |
| March 20 (Fri) | | Barbara Doshier, CELEST Science of Learning Seminar
<i>Mechanisms and modes of perceptual learning</i> |
| March 23 | | <i>Mid-term exam (1:00 – 3:00 PM)</i> Movie: <i>TBA</i> |
| March 30 | 8. | Physiology, psychology, and memory models. |
| April 6 | 9. | Content-addressable memories (CAM), active network design |
| April 13 | 10. | Liapunov functions, Cohen-Grossberg theorem |
| April 23 (Thurs) | 11. | Three-layer feedforward networks: Theory and mathematical foundations |
| April 27 | 12. | Synapses, signal functions, distributed vs. winner-take-all coding
<i>Course evaluations.</i> |
| May 11 | | <i>Final exam (1:00 - 3:30 PM)</i> <i>Class party</i> |
| extra | 13. | Invariance, spatial preprocessing, oscillations, temporal order information (TOI) |
| extra | 14. | Wavelets, genetic algorithms, fuzzy set theory |

CN550: TEXTBOOKS & READINGS

Most readings for CN550 are recent and classical research articles, reviews, and book chapters.

Many of the reference books and journals listed below are available in the CNS Library:

<http://www.cns.bu.edu/library/>

Please send me recommendations for additional books you would like to have in the Library.

The following books are sold in the Bookstore as CN550 texts.

Required:

Duda, Richard O., Hart, Peter E., & Stork, David (2001) *Pattern Classification*. Second Edition. New York: Wiley.

<http://rii.ricoh.com/~stork/DHS.html>

Recommended:

Schacter, Daniel L. (1996) *Searching for Memory: The Brain, the Mind, and the Past*.

New York: Basic Books. (paper)

Reprints from these texts are not posted on the web, neither are extensive readings from primary texts such as Levine (2000) or Kandel, Schwartz, & Jessell (2000), or some classical (and readily available) papers.

CN550 REFERENCES BOOKS

Amari, Shin-Ichi, & Arbib, Michael A. (Eds.) (1982) *Competition and Cooperation in Neural Nets*. Lecture Notes in Biomathematics, **45**. Berlin: Springer-Verlag.

Neural Information Processing Systems, Proceedings of the NIPS Conferences, New York: American Institute of Physics.

Anderson, James A., & Rosenfeld, Edward (Eds.) (1988) *Neurocomputing: Foundations of Research*. Cambridge, Mass.: MIT Press.

Anderson, James. A., Pellionisz, Andras, & Rosenfeld, Edward (Eds.) (1990) *Neurocomputing: Directions for Research*. Cambridge, Mass.: MIT Press.

Annual Review of Neuroscience.

Arbib, Michael A. (2003) *The Handbook of Brain Theory and Neural Networks, Second Edition*. Cambridge, Mass.: MIT Press.

Baddeley, Alan, Conway, Martin A., & Aggleton, John P. (Eds.) (2002) *Episodic Memory: New Directions in Research*. Oxford, UK: Oxford University Press.

Barabási, Albert-László (2003) *Linked: How Everything is Connected to Everything Else and What It Means*. Cambridge, MA: Perseus Publishing.

Bartlett, Frederic C. (1932) *Remembering: A Study in Experimental and Social Psychology*. Cambridge: Cambridge University Press.

- Bear, Mark F., Connors, Barry W., & Paradiso, Michael A. (1996) *Neuroscience: Exploring the Brain*. Baltimore: Williams & Wilkins.
- Bergson, Henri (1896) *Matter and Memory*. (Translated by N.M. Paul & W.S. Palmer, 1912). Zone Books.
- Bilder, Robert M., & LeFever, F. Frank (Eds.) (1998) *Neuroscience of the Mind on the Centennial of Freud's Project for a Scientific Psychology*. Annals of the New York Academy of Sciences, Vol. 843. New York: The New York Academy of Sciences.
- Bishop, Christopher M. (2006) *Pattern Recognition and Machine Learning*. Springer.
- Blanchard, Paul, Devaney, Robert L., & Hall, Glen R. (2002) *Differential Equations*. Second Edition. Pacific Grove, CA: Brooks/Cole.
- Carpenter, Gail A., & Grossberg, Stephen (Eds.) (1991) *Pattern Recognition by Self-Organizing Neural Networks*. Cambridge, Mass.: MIT Press.
- Cohen, Neal J., & Eichenbaum, Howard (1993) *Memory, Amnesia, and the Hippocampal System*. Cambridge, Mass.: MIT Press.
- Cooper, Jack R., Bloom, Floyd E., & Roth, Robert H. (2002) *The Biochemical Basis of Neuropharmacology*, Eighth Edition. New York: Oxford University Press.
- Colomb, Gregory G. (1990) *Style: Toward Clarity and Grace*. Chicago: The University of Chicago Press.
- Dasarathy, Belur V. (1991) *Nearest Neighbor (NN) Norms: NN Pattern Classification Techniques*. Washington, DC: IEEE Computer Society Press.
- Daubechies, Ingrid (1992) *Ten Lectures on Wavelets*. Philadelphia: Society for Industrial and Applied Mathematics.
- Davis, Paul W., & Moler, Cleve B. (1999) *Differential Equations: Modeling with Matlab*. Prentice Hall.
- Dodwell, Peter C., & Caelli, Terry (Eds.) (1984) *Figural Synthesis*. Hillsdale, NJ: Lawrence Erlbaum Asso.
- Dörner, Dietrich (1996) *The Logic of Failure: Recognizing and Avoiding Error in Complex Situations*. New York: Metropolitan Books. (Translated by Rita Kimber and Robert Kimber)
- Duda, Richard O., & Hart, Peter E. (1973) *Pattern Classification and Scene Analysis*. New York: Wiley.
- Duda, Richard O., Hart, Peter E., & Stork, David (2001) *Pattern Classification*. Second Edition. New York: Wiley.
- Ebbinghaus, Hermann (1885) *Memory : A Contribution to Experimental Psychology*. New York: Dover Publications, Inc. (1964 edition).
- Edelstein-Keshet, Leah (1988) *Mathematical Models in Biology*. SIAM Classics in Applied Mathematics, vol. 46.
- Freud, Sigmund (1886-1899) *Project for a Scientific Psychology*. In James Strachey (Ed.) *The Complete Psychological Works of Sigmund Freud*, Vol. I, London: Hogarth Press.
- Gazzaniga, Michael S. (Ed.) (2004) *The Cognitive Neurosciences III*. Third Edition. Cambridge, Mass.: MIT Press.
- Gladwell, Malcolm (2005). *Blink: The Power of Thinking Without Thinking*. New York: Little, Brown and Company.
- Griffith, N. & Todd, P.M. (Eds.) (1999) *Musical Networks*. MIT Press.
- Grossberg, Stephen (1982) *Studies of Mind and Brain: Neural Principles of Learning, Perception, Development, Cognition, and Motor Control*. Boston: Reidel / Kluwer Publ.
- Grossberg, Stephen (Ed.) (1987) *The Adaptive Brain I: Cognition, Learning, Reinforcement, and Rhythm*. Amsterdam: North-Holland.
- Grossberg, Stephen (Ed.) (1987) *The Adaptive Brain II: Vision, Speech, Language, and Motor Control*. Amsterdam: North-Holland.
- Grossberg, Stephen (Ed.) (1988) *Neural Networks and Natural Intelligence*. Cambridge, Mass.: MIT Press.
- Hawkins, Jeff (2004) *On Intelligence*. New York: Henry Holt.
- Haykin, Simon (1999) *Neural Networks - A Comprehensive Foundation*. Second Edition. Upper Saddle River, NJ: Prentice Hall.
- Hertz, John A., Krogh, Anders, & Palmer, Richard G. (1991) *Introduction to the Theory of Neural Computation*. Reading, Mass.: Addison-Wesley Publ.
- Hirsch, Morris W., Smale, Stephen, & Devaney, Robert L. (2004) *Differential Equations, Dynamical Systems, and an Introduction to Chaos*. Second Edition. Amsterdam: Academic Press.
- Hofstadter, Douglas R. (1979) *Gödel, Escher, Bach: an Eternal Golden Braid*. New York: Random House.
- Kandel, Eric R., Schwartz, James H., & Jessell, Thomas P. (Eds.) (2000) *Principles of Neural Science*. 4th Edition. New York: McGraw-Hill.
- Kanerva, Pentti (1988) *Sparse Distributed Memory*. Cambridge, Mass.: MIT Press.
- Kohonen, Teuvo (1988) *Self-Organization and Associative Memory*, Second Edition. Springer Series in Information Sciences, Vol. 8. Berlin: Springer-Verlag.

- Kosko, Bart (1992) *Neural Networks and Fuzzy Systems*. Englewood Cliffs, NJ: Prentice Hall.
- Kosslyn, Stephen M., & Andersen, Richard A. (Eds.) (1992) *Frontiers in Cognitive Neuroscience*. Cambridge, Mass.: MIT Press.
- Lem, Stanislaw (1967) *His Master's Voice*. Harvest Books. (Translated from Polish by Michael Kandel.)
- Levine, Daniel S. (2000) *Introduction to Neural and Cognitive Modeling*. Hillsdale, NJ: Lawrence Erlbaum Associates, 2nd Edition.
- Lin, C.C., & Segel, Lee A. (1974) *Mathematics Applied to Deterministic Problems in the Natural Sciences*. New York: Macmillan Publ.
- McGaugh, James L., Weinberger, Norman M., & Lynch, Gary (Eds.) (1995). *Brain and Memory: Modulation and Mediation of Neuroplasticity*. New York: Oxford University Press.
- Mead, Carver (1989) *Analog VLSI and Neural Systems*. Reading, Mass.: Addison-Wesley Publ.
- Meyer, Yves (Translated and revised by Robert D. Ryan) (1993) *Wavelets: Algorithms and Applications*. Philadelphia: Society for Industrial and Applied Mathematics.
- Minsky, Marvin, & Papert, Seymour (1969) *Perceptrons*. Cambridge, Mass.: MIT Press.
- Neopolitan, Richard E. (2004) *Learning Bayesian Networks*. Upper Saddle River, NJ: Prentice Hall.
- Nicholls, John G., Martin, A. Robert, Wallace, Bruce G., & Fuchs, Paul A., (2001) *From Neuron to Brain*. Fourth Edition. Sunderland, Mass.: Sinauer Asso.
- Penrose, Roger (1989) *The Emperor's New Mind*. New York: Oxford University Press.
- Ramachandran, V.S. (2004) *A Brief Tour of Human Consciousness: From Impostor Poodles to Purple Numbers*. New York: Pi Press.
- Reason, James (1990) *Human Error*. Cambridge, UK: Cambridge University Press.
- Rudin, Walter (1964) *Principles of Mathematical Analysis*. New York: McGraw-Hill.
- Rumelhart, David E., McClelland, James L., and the PDP Research Group (1986) *Parallel Distributed Processing: Explorations in the Microstructure of Cognition*, Vols. 1 & 2. Cambridge, Mass.: MIT Press.
- Schacter, Daniel L. (1996) *Searching for Memory: The Brain, the Mind, and the Past*. New York: Basic Books.
- Schacter, Daniel L. (2001) *The Seven Sins of Memory: How the Mind Forgets and Remembers*. Houghton Mifflin Co..
- Schwartz, Eric (Ed.) (1990) *Computational Neuroscience*. Cambridge, Mass.: MIT Press.
- Scott, Alwyn (1995) *Stairways to the Mind: The Controversial New Science of Consciousness*. New York: Springer-Verlag - Copernicus.
- Searle, John R. (1992) *The Rediscovery of the Mind*. Cambridge, Mass.: MIT Press.
- Shaw, Gordon L., McGaugh, James L., & Rose, Steven P.R. (Eds.) (1990) *Neurobiology of Learning and Memory - Reprint Volume*. World Scientific Advanced Series in Neuroscience Vol. 2, Teaneck, NJ: World Scientific Publ.
- Smith, Edward E., & Medin, Douglas L. (1981) *Categories and Concepts*. Cambridge, Mass.: Harvard University Press.
- Smith, Murray (1993) *Neural Networks for Statistical Modeling*. New York: Van Nostrand Reinhold.
- Squire, Larry R. (1987) *Memory and Brain*. New York: Oxford University Press.
- Squire, Larry R., & Kandel, Eric R. (2000) *Memory: From Mind to Molecules*. New York: Owl Books.
- Squire, Larry R., & Schacter, D. (Eds.) (2002) *The Neuropsychology of Memory*, Third Edition. New York: Guilford Publications.
- Strang, Gilbert (1988) *Linear Algebra and its Applications*, Third Edition. New York: Harcourt Brace Jovanovich College Publishers.
- Strunk, William, Jr., & White, E.B. (1979) *Elements of Style*, Third Edition. New York: Macmillan Publ.
- Surowiecki, James (2004) *The Wisdom of Crowds: Why the Many Are Smarter Than the Few and How Collective Wisdom Shapes Business, Economies, Societies and Nations*. Doubleday.
- Tulving, Endel, & Craik, Fergus I.M. (Eds.) (2000) *The Oxford Handbook of Memory*. New York: Oxford University Press.
- Vapnik, Vladimir N. (1998) *Statistical Learning Theory*. New York: John Wiley.
- Witten, Ian H., & Frank, Eibe (2005) *Data Mining: Practical Machine Learning Tools and Techniques, Second Edition*. Amsterdam: Morgan Kaufmann.
- Xu, Rui, & Wunsch, Don (2008) *Clustering*. Wiley – IEEE Press.

1. Overview, history, philosophy, benchmark database studies

Course goals, topics, methods, assignments

Historical review of principal neural network modules for learning, pattern recognition, and associative memory

Class project: Comparative studies of supervised learning systems

Benchmark database studies

Overview, history, philosophy

Daugman, John G. (1990) Brain metaphor and brain theory. In Eric Schwartz (Ed.) *Computational Neuroscience*. Cambridge, Mass.: MIT Press. Chapter 2: pp. 9-18.

Borges, Jorge Luis (1942) Funes, the Memorious. In: *Ficciones* (translation), New York: Grove Press (1962), pp. 107-115 <http://en.wikipedia.org/wiki/Borges>

Henig, Robin Marantz (2004) The quest to forget. *The New York Times Magazine*, April 4, 2004, pp. 32-37.

Treffert, Darold A., and Christensen, Daniel D. (2005) Inside the mind of a savant. *Scientific American*, Dec., pp. 108-113.

Carpenter, Gail A. (1989) Neural network models for pattern recognition and associative memory. *Neural Networks*, **2**, 243-257.

McCulloch, Warren S., & Pitts, Walter (1943) A logical calculus of the ideas immanent in nervous activity. *Bulletin of Mathematical Biophysics*, **5**, 115-133.

Bower, Gordon H. (2000) A brief history of memory research. In Endel Tulving & Fergus I.M. Craik, (Eds.) *The Oxford Handbook of Memory*. New York: Oxford University Press, Chapter 1, pp. 3-32.

Smith, Edward E., & Medin, Douglas L. (1981) *Categories and Concepts*. Cambridge, Mass.: Harvard University Press. Chapters 1-2, pp. 1-21.

Grossberg, Stephen (1982) *Studies of Mind and Brain*. Boston: Reidel / Kluwer Publ. - Preface, Introduction, and prefaces of chapters 1-13.

Memory extremes: Audio & video

<http://www.npr.org/templates/story/story.php?storyId=5352811>

Unique memory lets woman replay life like a movie

Morning Edition, April 20, 2006 · Neurobiologist James McGaugh, one of the world's experts on human memory, says that a woman he calls AJ has a one-of-a-kind memory. In an interview with NPR, she talks about what life is like for someone who can remember things she's done and news events from almost every day of her life for the past 25 years. Her life is like a split-screen movie, with the past running almost as vividly as the present.

Clive Wearing: Living without memory YouTube (BBC – The Mind): [Pt2a](#) [Pt2b](#) [Pt2c](#) [Pt2d](#)

http://en.wikipedia.org/wiki/Clive_Wearing

Wikipedia: Clive Alex Wearing (born 1938) is a [British musicologist](#), [conductor](#), and [keyboardist](#) suffering from an acute and long lasting case of [anterograde amnesia](#). Specifically, this means he lacks the ability to form new memories, dubbed the "memento" syndrome by laypeople and the media, after a film based on the subject [of the same name](#).

Class project databases

CN710 wikis – including financial database reports and results

Fall 2008 <http://cns.bu.edu/cn710/Fall2008/>
Fall 2007 <http://cns.bu.edu/cn710/Fall2007/>
Fall 2006 <http://cns.bu.edu/cn710/Fall2006/pmwiki.php?n=Main.HomePage>
Spring 2006 <http://cns.bu.edu/cn710/Spring2006/pmwiki.php?n=Main.HomePage>

Financial prediction

Versace, Massimiliano, Bhatt, Rushi, Hinds, Oliver, & Shiffer, Mark (2004) Predicting the exchange traded fund DIA with a combination of genetic algorithms and neural networks. *Expert Systems with Applications*, 27(3), 417-425. [This paper, by four CNS students, began with the CN550 class project.]

UCI Repository

Hettich, S., & Bay, S.D. (1999) The UCI KDD Archive. Irvine, CA: University of California, Department of Information and Computer Science. <http://www.ics.uci.edu/~mlearn/MLRepository.html>

Response plots

Lang, K.J., & Witbrock, M.J. (1989) Learning to tell two spirals apart. In David S. Touretzky, Geoffrey Hinton, & Terrence Sejnowski (Eds.) *Proceedings of the 1988 Connectionist Models Summer School*. San Mateo, Calif.: Morgan Kaufmann Publ. pp. 52-59.

Writing

Strunk, William, Jr., & White E.B. (1959-2000) *The Elements of Style*, Fourth Edition. Needham Heights, MA: Allyn & Bacon.

APA style: <http://www.apastyle.org/> , <http://www.bridgewater.edu/WritingCenter/manual/APAformat.htm>

2. Supervised learning methods: Memory-based algorithms (KNN), model-independent supervised learning methods (validation & cross-validation, c-index, ROC curves, resampling, combining classifiers, component analysis), statistical pattern recognition

Memory-based algorithms: K-nearest neighbors (K-NN)

Approaching supervised learning problems fairly and systematically

Training, testing, validation, and cross-validation

ROC curves and the c-index

Resampling: bootstrapping, boosting, bagging

Combining systems: mixing models and voting

Data preparation: component analysis

Brief introduction to statistical pattern recognition and Bayesian estimation

Memory-based algorithms <http://en.wikipedia.org/wiki/KNN>

Duda, Richard O., Hart, Peter E., & Stork, David (2001) *Pattern Classification*. Second Edition. New York: Wiley. Section 4.1-4.6: Nonparametric techniques, pp. 161-192

Training, testing, validation, and cross-validation

Duda, Richard O., Hart, Peter E., & Stork, David (2001) *Pattern Classification*. Second Edition. New York: Wiley. Section 9.6.2: Cross-validation, pp. 483-485.

C-index and ROC curves http://en.wikipedia.org/wiki/Roc_curve

Duda, Richard O., Hart, Peter E., & Stork, David (2001) *Pattern Classification*. Second Edition. New York: Wiley. Section 2.8.3: Signal detection theory and operating characteristics, pp. 48-51.

Resampling: Bootstrapping, boosting, bagging http://en.wikipedia.org/wiki/Resampling_%28statistics%29

Duda, Richard O., Hart, Peter E., & Stork, David (2001) *Pattern Classification*. Second Edition. New York: Wiley. Section 9.4: Resampling for estimating statistics, pp. 471-475.

9.4.1 Jackknife http://en.wikipedia.org/wiki/Resampling_%28statistics%29#Jackknife

9.4.2 Bootstrap http://en.wikipedia.org/wiki/Bootstrapping_%28statistics%29

Duda, Richard O., Hart, Peter E., & Stork, David (2001) *Pattern Classification*. Second Edition. New York: Wiley. Section 9.5: Resampling for classifier design, pp. 475-482

9.5.1 Bagging http://en.wikipedia.org/wiki/Bootstrap_aggregating

9.5.2 Boosting <http://en.wikipedia.org/wiki/Boosting>

[CONTINUED ON THE FOLLOWING PAGE]

Mixing models and voting

Duda, Richard O., Hart, Peter E., & Stork, David (2001) *Pattern Classification*. Second Edition. New York: Wiley. Section 9.7: Combining classifiers, pp. 495-499.

Carpenter, Gail A., Grossberg, Stephen, Markuzon, Natalya, Reynolds, John H., & Rosen, David B. (1992) Fuzzy ARTMAP: A neural network architecture for incremental supervised learning of analog multidimensional maps. *IEEE Transactions on Neural Networks*, **3**, 698-713.

Component analysis

Duda, Richard O., Hart, Peter E., & Stork, David (2001) *Pattern Classification*. Second Edition. New York: Wiley. Section 3.8: Component analysis and discriminants, pp. 114-124

- 3.8.1 Principal component analysis (PCA) http://en.wikipedia.org/wiki/Principal_Component_Analysis
- 3.8.2 Fisher linear discriminant http://en.wikipedia.org/wiki/Fisher_linear_discriminant
- 3.8.3 Multiple discriminant analysis

Maximum-likelihood and Bayesian parameter estimation http://en.wikipedia.org/wiki/Thomas_Bayes

Duda, Richard O., Hart, Peter E., & Stork, David (2001) *Pattern Classification*. Second Edition. New York: Wiley, pp. 84-97

- 3.1 Introduction
- 3.2 Maximum-likelihood estimation (MLE) http://en.wikipedia.org/wiki/Maximum_likelihood
- 3.3 Bayesian estimation http://en.wikipedia.org/wiki/Bayes%27_theorem
- 3.4 Bayesian parameter estimation: Gaussian case

3. Unsupervised learning: Clustering (leader, K-means), competitive learning, ART

Clustering algorithms: Leader clustering and K-means clustering

Norms and metrics

Competitive learning

Adaptive resonance theory - 1970s

ART 1: Binary pattern learning

ART 2-A: A fast, algorithmic version of ART 2

Freud's neural networks

Clustering http://en.wikipedia.org/wiki/Data_clustering

Duda, Richard O., Hart, Peter E., & Stork, David (2001) *Pattern Classification*. Second Edition. New York: Wiley.
Section 10.4.3: k-means clustering, pp. 526-528 <http://en.wikipedia.org/wiki/K-means>

Competitive learning

Levine, Daniel S. (2000) *Introduction to Neural and Cognitive Modeling*. Hillsdale, NJ: Lawrence Erlbaum Associates, 2nd Edition.

Chapter 4: Competition, lateral inhibition, and short-term memory, pp. 95-154.

Chapter 6: Coding and categorization, pp. 198-279.

Malsburg, Christoph von der (1973) Self-organization of orientation sensitive cells in the striate cortex. *Kybernetik*, **14**, 85-100.

Grossberg, Stephen (1976) Adaptive pattern classification and universal recoding, I: Parallel development and coding of neural feature detectors. *Biological Cybernetics*, **23**, 121-134.

Adaptive Resonance Theory - 1970s

Grossberg, Stephen (1976) Adaptive pattern classification and universal recoding, II: Feedback, expectation, olfaction, and illusions. *Biological Cybernetics*, **23**, 187-202.

ART 1

Carpenter, Gail A., & Grossberg, Stephen (1987) A massively parallel architecture for a self-organizing neural pattern recognition machine. *Computer Vision, Graphics, and Image Processing*, **37**, 54-115.

Moore, Barbara (1989) ART 1 and pattern clustering. In David S. Touretzky, Geoffrey Hinton, & Terrence Sejnowski (Eds.) *Proceedings of the 1988 Connectionist Models Summer School*. San Mateo, Calif.: Morgan Kaufmann Publishers. pp. 174-185.

ART 2-A algorithm

Carpenter, Gail A., Grossberg, Stephen, & Rosen, David B. (1991) ART 2-A: An Adaptive Resonance algorithm for rapid category learning and recognition. *Neural Networks*, **4**, 493-504.

[CONTINUED ON THE FOLLOWING PAGE]

Freud's neural networks <http://en.wikipedia.org/wiki/Freud>

Freud, Sigmund (1886-1899) *Project for a Scientific Psychology*. pp. 322-325. (1900) *The Interpretation of Dreams*. Introduction by James Strachey (Editor and translator). New York: Avon Books (1965).

Tucker, Don M., & Luu, Phan (1998) Cathexis revisited: Corticolimbic resonance and the adaptive control of memory. In Robert M. Bilder & F. Frank LeFever (Eds.) *Neuroscience of the Mind on the Centennial of Freud's Project for a Scientific Psychology*. Annals of the New York Academy of Sciences, Vol. 843. New York: The New York Academy of Sciences. pp. 134-152.

4. Dimensional analysis, competitive networks, phase plane analysis

Dimensional analysis

Dynamics of on-center off-surround shunting competitive networks

Phase plane analysis of competitive networks

Dimensional analysis http://en.wikipedia.org/wiki/Dimensional_analysis

Lin, C.C., & Segel, L.A. (1974) *Mathematics Applied to Deterministic Problems in the Natural Sciences*. New York: Macmillan.

Chapter 6: Simplification, dimensional analysis, and scaling, pp. 185-224

Edelstein-Keshet, Leah (1988) *Mathematical Models in Biology*. SIAM Classics in Applied Mathematics, vol. 46.

Section 4.3: Formulating a model

Section 4.4: Saturating nutrient consumption rate

Section 4.5: Dimensional analysis of the equations

Phase plane analysis http://en.wikipedia.org/wiki/Phase_plane

Edelstein-Keshet, Leah (1988) *Mathematical Models in Biology*. SIAM Classics in Applied Mathematics, vol. 46.

Sections 5.2-5.9: Phase-plane methods and qualitative solutions , pp. 171-193.

Blanchard, Paul, Devaney, Robert L., & Hall, Glen R. (2002) *Differential Equations*. Second Edition. Pacific Grove, CA: Brooks/Cole. [+ solutions]

Boston University Ordinary Differential Equations Project: <http://math.bu.edu/odes/>

Section 3.3: Phase planes for linear systems with real eigenvalues, pp. 266-282.

Section 5.2: Qualitative analysis, pp. 457-470.

Section 5.3: Hamiltonian systems, pp. 470-488. http://en.wikipedia.org/wiki/Hamiltonian_mechanics

Section 5.4: Dissipative systems, pp. 488-510. <http://en.wikipedia.org/wiki/Dissipative>

5. ARTMAP

Fuzzy ART: Generalized ART 1, for analog inputs, using the city-block metric (L1 norm)

Supervised learning by ART systems

Binary ARTMAP

Analog fuzzy ARTMAP

Fuzzy ART

Carpenter, Gail A., Grossberg, Stephen, & Rosen, David B. (1991) Fuzzy ART: Fast stable learning and categorization of analog patterns by an Adaptive Resonance system. *Neural Networks*, **4**, 759-771.

Supervised learning by ARTMAP systems

Carpenter, Gail A., Grossberg, Stephen, & Reynolds, John H. (1991) ARTMAP: Supervised real-time learning and classification of nonstationary data by a self-organizing neural network. *Neural Networks*, **4**, 565-588.

Fuzzy ARTMAP and other ARTMAP networks

Carpenter, Gail A., Grossberg, Stephen, Markuzon, Natalya, Reynolds, John H., & Rosen, David B. (1992) Fuzzy ARTMAP: A neural network architecture for incremental supervised learning of analog multidimensional maps. *IEEE Transactions on Neural Networks*, **3**, 698-713.

Carpenter, Gail A. (2003). Default ARTMAP. *Proceedings of the International Joint Conference on Neural Networks (IJCNN'03)*, Portland, Oregon, 1396-1401.

Letter recognition & genetic algorithms

Frey, Peter W., & Slate, David J. (1991) Letter recognition using Holland-style adaptive classifiers. *Machine Learning*, **6**, 161-182.

Introduction to fuzzy set theory and fuzzy classification

http://en.wikipedia.org/wiki/Fuzzy_sets

http://en.wikipedia.org/wiki/Fuzzy_logic

Zadeh, Lotfi A. (1965) Fuzzy sets. *Information Control*, **8**, 338-353.

6. Associative memory networks: Back propagation, multi-layer perceptrons, radial basis functions, cascade-correlation, higher-order networks

Back propagation

Multi-layer perceptrons

(Local) minimization of cost functions

Radial basis functions (RBFs)

Cascade-correlation architecture

Higher order networks

The perceptron <http://en.wikipedia.org/wiki/Perceptron>

Rosenblatt, F. (1958) The perceptron: A probabilistic model for information storage and organization in the brain. *Psychological Review*, **65**, 386-408.

Duda, Richard O., Hart, Peter E., & Stork, David (2001) *Pattern Classification*. Second Edition. New York: Wiley. Sections 5.1-5.8: Linear discriminant functions, pp. 215-249

http://en.wikipedia.org/wiki/Linear_discriminant_analysis

Back propagation and cost functions http://en.wikipedia.org/wiki/Multilayer_perceptron
http://en.wikipedia.org/wiki/Back_propagation http://en.wikipedia.org/wiki/Cost_function

Duda, Richard O., Hart, Peter E., & Stork, David (2001) *Pattern Classification*. Second Edition. New York: Wiley. Sections 6.1-6.8: Multilayer neural networks, pp. 282-318.

Rumelhart, David E., Hinton, Geoffrey E., & Williams, Ronald J. (1986) Learning internal representations by error propagation. In David E. Rumelhart & James L. McClelland (Eds.), *Parallel Distributed Processing: Explorations in the Microstructure of Cognition I*. Cambridge, Mass.: MIT Press. pp. 318-362.

Radial basis functions (RBFs) http://en.wikipedia.org/wiki/Radial_basis_function

Moody, John, & Darken, Christian J. (1989) Fast learning in networks of locally-tuned processing units. *Neural Computation*, **1**, 281-294.

Lowe, David (2003) Radial basis function networks. In Arbib, Michael A. (2003) *The Handbook of Brain Theory and Neural Networks, Second Edition*. Cambridge, Mass.: MIT Press. pp. 937-940.

Duda, Richard O., Hart, Peter E., & Stork, David (2001) *Pattern Classification*. Second Edition. New York: Wiley. Section 6.10.1: Radial basis function networks (RBFs), pp. 324-325

Cascade-correlation network http://en.wikipedia.org/wiki/Cascade_correlation

Fahlman, Scott E., & Lebiere, Christian (1990) The cascade-correlation learning architecture. In David S. Touretzky (Ed.) *Neural Information Processing Systems 2*, Proceedings of the NIPS Conference, Denver, 1989, San Mateo, Calif.: Morgan Kaufmann Publishers. pp. 524-532. <http://www.cs.iastate.edu/~honavar/fahlman.pdf>

Duda, Richard O., Hart, Peter E., & Stork, David (2001) *Pattern Classification*. Second Edition. New York: Wiley. Section 6.10.6: Cascade-correlation, pp. 329-330.

Higher order networks

Giles, C. Lee, & Maxwell, Thomas (1987) Learning, invariance, and generalization in high-order neural networks. *Applied Optics*, **26**, 4972-4978.

7. Support vector machines

Support vector machines (SVMs)
Constrained optimization
Lagrange multipliers
Structural risk minimization
VC dimension

Constrained optimization and Lagrange multipliers

Vapnik, Vladimir N. (1998) *Statistical Learning Theory*. New York: John Wiley. Section 9.5: Three theorems of optimization theory, pp. 390-394. http://en.wikipedia.org/wiki/Optimization_%28mathematics%29

Strang, Gilbert (1988) *Linear Algebra and its Applications*, Third Edition. New York: Harcourt Brace Jovanovich College Publishers. Section 8.3: The theory of duality, pp. 412-423.
http://en.wikipedia.org/wiki/Dual_space

Duda, Richard O., Hart, Peter E., & Stork, David (2001) *Pattern Classification*. Second Edition. New York: Wiley. Section A.3: Lagrange optimization, p. 610.
http://en.wikipedia.org/wiki/Constrained_Optimization_and_lagrange_Multipliers

Bishop, Christopher M. (2006) *Pattern Recognition and Machine Learning*. Springer. Appendix E: Lagrange multipliers, pp. 707-710.

Support vector machines http://en.wikipedia.org/wiki/Support_vector_machines

Duda, Richard O., Hart, Peter E., & Stork, David (2001) *Pattern Classification*. Second Edition. New York: Wiley. Section 5.11: Support vector machines, pp. 259-265

Vapnik, Vladimir N. (1998) *Statistical Learning Theory*. New York: John Wiley. Chapter 10: The support vector method for estimating indicator functions, pp. 401 - 441.

V-C dimension

Bartlett, Peter L., & Maass, Wolfgang (2003) Vapnik-Chervonenkis dimension of neural nets. In Arbib, Michael A. (2003) *The Handbook of Brain Theory and Neural Networks, Second Edition*. Cambridge, Mass.: MIT Press. pp. 1188-1192.

Additional SVM resources

CN710 Readings re: SVM, especially the [Bennett and Campbell paper](#) .

8. Physiology, psychology, and memory models

Neural substrates of memory

Cortical organization

Neuropsychology of memory and amnesia

Synaptic modification

Redistribution of synaptic efficacy: What is a weight?

Schacter, Daniel L. (1996) *Searching for Memory: The Brain, the Mind, and the Past*. New York: Basic Books.

NOTE: This is an optional CN550 textbook. Try at least to skim it, reading sections that interest you.

Neural substrates of memory

http://en.wikipedia.org/wiki/Cerebral_cortex

Levine, Daniel S. (2000) *Introduction to Neural and Cognitive Modeling*. Hillsdale, NJ: Lawrence Erlbaum Associates, 2nd Edition.

Appendix 1: Basic Facts of Neurobiology, pp. 375-395.

<http://en.wikipedia.org/wiki/Neuron>

Bear, Mark F., Connors, Barry W., & Paradiso, Michael A. (1996) *Neuroscience: Exploring the Brain*. Baltimore: Williams & Wilkins. Chapter 19: Memory systems, pp. 514-545.

Neuropsychology of memory and amnesia

Clive Wearing: Living without memory YouTube (BBC – The Mind): [Pt2a](#) [Pt2b](#) [Pt2c](#) [Pt2d](#)

Hall, Stephen S. (1998) Our memories, our selves. *The New York Times Magazine* (February 15, 1998), 26-33, 49, 56-57.

Corkin, Suzanne (2002) What's new with the amnesic patient H.M.? *Nature Reviews - Neuroscience*, **3**, 153-160.

http://en.wikipedia.org/wiki/HM_%28patient%29

Freedman, David J., Riesenhuber, Maximilian, Poggio, Tomaso, & Miller, Earl K. (2003) A comparison of primate prefrontal and inferior temporal cortices during visual categorization. *Journal of Neuroscience*, **23**, 5235-5246.

Holt, Jim (2008) [Numbers Guy](#) The New Yorker, March 3, 2008, 42-47.

Synaptic modification http://en.wikipedia.org/wiki/Long-term_potentiation

Kandel, Eric R., Schwartz, James H., & Jessell, Thomas P. (Eds.) (2000) *Principles of Neural Science*. 4th Edition. New York: McGraw-Hill.

Chapter 63: Eric R. Kandel. Cellular mechanisms of learning and the biological basis of individuality. pp. 1247-1279.

Malenka, Robert C., & Nicoll, Roger A. (1999) Long-term potentiation - a decade of progress? *Science*, **285**, 1870-1874. Robert Malenka podcast – Synaptic plasticity: Multiple mechanisms and functions

<http://videocast.nih.gov/podcast.asp?13746>

Redistribution of synaptic efficacy

Markram, Henry, & Tsodyks, Misha (1996) Redistribution of synaptic efficacy between neocortical pyramidal neurons. *Nature*, **382**, 807-810.

Carpenter, Gail A. (2001) Neural network models of learning and memory: leading questions and an emerging framework. *Trends in Cognitive Sciences*, **5**, 114-118.

9. Content-addressable memories (CAM), active network design

Content-Addressable Memory (CAM)

Net STM activity of statistical ensembles

Nonspecific modulation and active regulation of STM

Principled construction of neural network models

Competitive networks

Grossberg, Stephen (1973) Contour enhancement, short term memory, and constancies in reverberating neural networks. *Studies in Applied Mathematics*, **LII**, 213-257.

STM system design

Ellias, Samuel A., & Grossberg, Stephen (1975) Pattern formation, contrast control, and oscillations in the short term memory of shunting on-center off-surround networks.

Biological Cybernetics, **20**, 69-98.

10. Liapunov functions, Cohen-Grossberg theorem

Liapunov functions and the LaSalle invariance principle
The Cohen-Grossberg theorem

http://en.wikipedia.org/wiki/Lyapunov_function

<http://mathworld.wolfram.com/LyapunovFunction.html>

The Cohen-Grossberg theorem

Cohen, Michael A., & Grossberg, Stephen (1983) Absolute stability of global pattern formation and parallel memory storage by competitive neural networks. *IEEE Transactions on Systems, Man, and Cybernetics*, **SMC-13**, 815-826.

Grossberg, Stephen (1988) Nonlinear neural networks: Principles, mechanisms, and architectures. *Neural Networks*, **1**, 17-61.

Section 9 - Content-addressable memory storage: a general STM model and Liapunov method, pp. 24 - 30.

Additional resource

Brauer, Fred, & Nohel, John (1969). The qualitative theory of differential equations. W.A. Benjamin.
Sections 5.1 and 5.2.

11. Three-layer feedforward networks: Theory and mathematical foundations

Mappings by three-layer feedforward networks: mathematical analysis

Mapping by 3-layer feedforward networks

Hecht-Nielsen, Robert (1987) Kolmogorov's mapping neural network existence theorem. *Proceedings, IEEE First International Conference on Neural Networks*, San Diego: SOS Press. pp. III:11-14.

http://en.wikipedia.org/wiki/Hilbert%27s_problems

http://en.wikipedia.org/wiki/Hilbert%27s_thirteenth_problem

Duda, Richard O., Hart, Peter E., & Stork, David (2001) *Pattern Classification*. Second Edition. New York: Wiley. Section 6.2.2: Expressive power of multilayer networks, pp. 287-288

Hornik, Kurt, Stinchcombe, Maxwell, & White, Halbert (1989) Multi-layer feedforward networks are universal approximators. *Neural Networks*, **2**, 359-366.

Cybenko, George (1989) Approximation by superpositions of a sigmoidal function. *Mathematics of Control, Signals, and Systems*, **2**, 303-314.

Reference book: Mathematical analysis

Rudin, Walter (1964) *Principles of Mathematical Analysis*. New York: McGraw-Hill.

12. Synapses, signal functions, distributed learning

Neurobiology of chemical synapses, neuromodulators, and short-term synaptic plasticity

Retrograde messengers

ART 3: Chemical transmitters, retrograde messengers, and neuromodulators for implementing parallel memory search

Distributed outstar learning and rules of synaptic transmission

Physiology of chemical synapses

Kandel, Eric R., Schwartz, James H., & Jessell, Thomas P. (Eds.) (2000) *Principles of Neural Science*. 4th Edition. New York: McGraw-Hill, pp. 175-186.

Chapter 10: Eric R. Kandel & Steven A. Siegelbaum. Overview of synaptic transmission.

<http://en.wikipedia.org/wiki/Synapse>

Zucker, Robert S. (1989) Short-term synaptic plasticity. *Annual Review of Neuroscience*, **12**, pp. 13-31.

Retrograde messengers and ART 3 http://en.wikipedia.org/wiki/Retrograde_signaling_in_LTP

Kiss, János P., & Vizi, E. Sylvester (2001) Nitric oxide: a novel link between synaptic and nonsynaptic transmission. *Trends in Neurosciences*, **24**, 211-215.

http://en.wikipedia.org/wiki/Nitric_oxide

Carpenter, Gail A., & Grossberg, Stephen (1990) ART 3: Hierarchical search using chemical transmitters in self-organizing pattern recognition architectures. *Neural Networks*, **3**, 129-152.

Rules of synaptic transmission and distributed outstar learning

Carpenter, Gail A. (1994) A distributed outstar network for spatial pattern learning. *Neural Networks*, **7**, 159-168.

13 [extra]. Invariance, spatial preprocessing, oscillations, temporal order information (TOI)

Invariant pattern recognition

Fourier analysis

Log-polar-Fourier filter

Singular solutions

Hopf bifurcation

Coding of temporal order information for event sequences

Free recall paradigm

Models of TOI: buffer, activation gradient, oscillations

Image transforms and preprocessing

http://en.wikipedia.org/wiki/Fourier_transform

http://en.wikipedia.org/wiki/Complex_logarithm

Cavanagh, Patrick (1984) Image transforms in the visual system. In Peter C. Dodwell & Terry Caelli (Eds.) *Figural Synthesis*. Hillsdale, NJ: Lawrence Erlbaum Associates. pp. 185-218.

Duda, Richard O., & Hart, Peter E. (1973) *Pattern Classification and Scene Analysis*. New York: Wiley.
Chapter 8: The spatial frequency domain, pp. 298-326

Oscillations in dynamical systems: Hopf bifurcation and singular solutions

Ellias, Samuel A., & Grossberg, Stephen (1975) Pattern formation, contrast control, and oscillations in the short term memory of shunting on-center off-surround networks. *Biological Cybernetics*, **20**, 69-98.

http://en.wikipedia.org/wiki/Hopf_bifurcation

Carpenter, Gail A. (1981) Normal and abnormal signal patterns in nerve cells. In Stephen Grossberg (Ed.), *Mathematical Psychology and Psychophysiology*, AMS/SIAM Symposium Series, **13**, pp. 49 -90.

http://en.wikipedia.org/wiki/Hodgkin-Huxley_model

Temporal order information (TOI)

Grossberg, Stephen (1978) A theory of human memory: Self-organization and performance of sensory-motor codes, maps, and plans. In *Progress in Theoretical Biology*, **5**, New York: Academic Press, pp. 233 - 374. Reprinted in *Studies of Mind and Brain*, Chapter 13, pp. 498-639.

Sections 1, 23-38 (From *Studies of Mind and Brain*, pp. 500-512, 564-591.)

Buffer models http://en.wikipedia.org/wiki/Short-term_memory

Atkinson, Richard C., & Shiffrin, Richard M. (1971) The control of short-term memory. *Scientific American*, 82-90. http://en.wikipedia.org/wiki/Atkinson-Shiffrin_memory_model

Grossberg, Stephen (1978) A theory of human memory: Self-organization and performance of sensory-motor codes, maps, and plans. In *Progress in Theoretical Biology*, **5**, New York: Academic Press, pp. 233-374.

Section 61: Automatic versus controlled information processing (SMB: pp. 626-631)

14 [extra]. Wavelets, genetic algorithms, fuzzy set theory

Wavelets

Introduction to genetic algorithms

Fuzzy set theory: definitions

Wavelets

<http://en.wikipedia.org/wiki/Wavelet>

Daubechies, Ingrid (1992) *Ten Lectures on Wavelets*. Philadelphia: Society for Industrial and Applied Mathematics.

Preliminaries and notation, pp. xi - xix.

Chapter 1: The what, why, and how of wavelets , pp. 1 - 16.

Daugman, John G. (2001) How iris recognition works. <http://www.CL.cam.ac.uk/users/jgd1000/>

<http://www.cl.cam.ac.uk/~jgd1000/csvt.pdf>

Introduction to genetic algorithms

http://en.wikipedia.org/wiki/Genetic_algorithm

Duda, Richard O., Hart, Peter E., & Stork, David (2001) *Pattern Classification*. Second Edition. New York: Wiley.

Section 7.5.1: Genetic algorithms, pp. 373-377.

Booker, L.B., Goldberg, D.E., & Holland, J.H. (1989) Classifier systems and genetic algorithms. *Artificial Intelligence*, **40**, 235-282.

Frey, Peter W., & Slate, David J. (1991) Letter recognition using Holland-style adaptive classifiers. *Machine Learning*, **6**, 161-182.

Introduction to fuzzy set theory and fuzzy classification

http://en.wikipedia.org/wiki/Fuzzy_sets

http://en.wikipedia.org/wiki/Fuzzy_logic

Zadeh, Lotfi A. (1965) Fuzzy sets. *Information Control*, **8**, 338-353.

Duda, Richard O., Hart, Peter E., & Stork, David (2001) *Pattern Classification*. Second Edition. New York: Wiley.

Section 4.7: Fuzzy classification, pp. 192-195.

Duda, Richard O., Hart, Peter E., & Stork, David (2001) *Pattern Classification*. Second Edition. New York: Wiley.

Section 10.4.4: Fuzzy k-means clustering, pp. 528-530.