ECEN 649 Pattern Recognition – Spring 2012
Class Meetings: TR 12:45PM–2:00PM
Room: ETB 1003
Credits: 3

Course Objective: This course aims to introduce the basic elements of Statistical Pattern Recognition, focusing on critical mathematical and statistical aspects underlying pattern recognition methods. After a brief review of probability theory, we will cover fundamental concepts of classification, such as Bayesian classification and classification consistency, followed by a study of several families of classification rules, error estimation, dimensionality reduction, and model selection, including an introduction to Vapnik-Chervonenkis theory. Performance will be assessed by means of one midterm exam and one final exam, four homework assignments, and one computer project.

Instructor: Ulisses M. Braga-Neto
Room: 237N
E-mail: `ulisses@ece.tamu.edu`
Office Hours: TR 2:15PM – 3:00PM (or by appointment)

Class Website: `http://www.ece.tamu.edu/~ulisses/ECEN649/index.html`

Primary Texts (on reserve at Evans Library):

1. L. Devroye, L. Gyorfi and G. Lugosi, *A Probabilistic Theory of Pattern Recognition*, Springer, 1996 (DGL).
2. R. Duda, P. Hart and D. Stork, *Pattern Classification*, 2nd ed., John Wiley & Sons, 2001 (DHS).
3. A. Webb, *Statistical Pattern Recognition*, 2nd ed., John Wiley & Sons, 2002 (Webb).

Supplementary Reading:

1. G.J. McLachlan, *Discriminant Analyis and Statistical Pattern Recognition*, John Wiley & Sons, 2004.
2. K. Fukunaga, *Statistical Pattern Recognition*, 2nd ed., Academic Press, 1990.
3. V. Vapnik, *Statistical Learning Theory*, John Wiley & Sons, 1998.
4. A.K. Jain, R.P.W. Duin, and J. Mao. Statistical pattern recognition: A review. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(1):4–37, 2000.
5. U.M. Braga-Neto and E.R. Dougherty. Classification. In *Genomic Signal Processing and Statistics*, EURASIP Book Series on Signal Processing and Communication. Hindawi Publishing Corporation, 2005.
6. U.M. Braga-Neto. Fads and Fallacies in the Name of Small-Sample Microarray Classification. *IEEE Signal Processing Magazine*, January 2007.

Course Topics:

1. Introduction to Pattern Recognition
2. Review of Probability (class notes)
    - Sample Space, Events, Definition of Probability
    - Borel-Cantelli Lemmas
    - Conditional Probability, Independence
    - Random Variables
    - Expectation, Variance, Correlation
    - Joint distributions
    - Conditional Expectation and Prediction
    - Multivariate Gaussian distribution
    - Random sequences and convergence
    - Laws of large numbers, Central Limit Theorem
3. Fundamental Concepts in Classification
    - Bayes Decision Theory (DHS 2.1–2.4,2.7; DGL 2.1–2.4)
    - Multivariate Gaussian Case (DHS 2.5–2.6)
    - Alternate Distance Measures (DGL 3.1,3.3–3.5,3.7–3.8; DHS 2.8.1–2.8.2)
    - Classification Consistency (DGL 2.5, 6.1-6.2)
4. Classification Rules
    - Parametric Discriminant Analysis (DHS 3.8.2; DGL 4.3,16.0,16.2; class notes)
    - Histogram Rule (DGL 6.3,6.4)
    - Nearest-Neighbor Rules (DGL 5.1–5.2, 5.4, 5.7–5.8; DHS 4.4–4.5)
    - Kernel-Based Methods (DGL 10.0–10.1,10.3; DHS 4.1–4.3)
    - Perceptrons (DHS 5.1–5.2, 5.4–5.7)
    - Support-Vector Machines (DHS 5.11; Webb 4.2.5, 5.4.0–5.4.2)
    - Neural Networks (DHS 6.1-6.3; DGL 30.1–30.2, 30.4(Thm 30.7), 30.5; Webb 6.2.1–6.2.4)
    - Decision Trees (DHS 8.1-8.3; DGL 20.0–20.1, 20.9, 20.11)
    - Discrete Classification (DGL 27.0–27.1; class notes)
5. Error Estimation (DGL 8.1–8.3, 8.5, 23.0–23.2, 24.0, 24.2–24.5; class notes)
6. Dimensionality Reduction (DGL 32.1, 32.4; Webb 9.1–9.3.2, 9.4.0, 9.4.2)
7. Model Selection and Vapnik-Chervonenkis Theory (DGL 8.4, 12.1, 12.4–12.6, 13.1–13.3, 14.3–14.4, 18.1, 30.4; DHS 6.3.3, 9.6.2, 9.6.8; Webb 11.1.1–11.1.2)

Grading:

- Homework: 4 assignments, ∼1 every 3 weeks (**25%**)
- Midterm: March 6 (during class) (**25%**)
- Final: May 4, 12:30–2:30 pm (**35%**)
- Computer Project (**15%**)

You are allowed to use the class notes and the references listed on this syllabus in the solution of the assignments, but no other external sources. One A4 sheet of paper with formulas (front and back) is allowed for the Midterm exam, whereas 2 sheets are allowed for the Final exam.

Americans with Disabilities Act (ADA) Policy Statement:

> The Americans with Disabilities Act (ADA) is a federal antidiscrimination statute that provides comprehensive civil rights protection for persons with disabilities. Among other things, this legislation requires that all students with disabilities be guaranteed a learning environment that provides for reasonable accommodation of their disabilities. If you believe you have a disability requiring an accommodation, please contact the Department of Student Life, Services for Students with Disabilities in Cain Hall, Rm. B118, or call 845-1637.

Aggie Code of Honor:

"An Aggie does not lie, cheat, or steal, or tolerate those who do."

For detailed information on Texas A&M academic integrity policies, please visit `http://www.tamu.edu/aggiehonor/know.html`