Instructor:
Professor J.R. Green, P.Eng., room 6203CB, e-mail jrgreen@sce.carleton.ca (Please send from your connect account and include “SYSC5405/BIOM5405” in the subject line to ensure a response.)

Office hours: By appointment. See course CULearn site for additional hours.

Instructional Hours per Week:
3 lecture hours: Mondays & Wednesdays 13:05-14:25. Check Carleton Central for location.

Prerequisites:
Undergraduate introductory probability and statistics. (Plus graduate standing, permission of the instructor or the department.)

Primary Textbook:

Other References:
5. Current literature describing applications of pattern classification.
6. Freely available MATLAB/JAVA/C++ implementations of pattern classification methods (e.g. Netlab and WEKA).
7. Additional materials may be made available on the course website.

Items 1, 2, and 3 are on reserve at the CU Library.

Web Page:
All content is posted on the CULearn system. University of Ottawa students need to complete a form and submit it to the FGPA to gain access to CULearn for the term. Students are required to check CULearn often for course updates. Supplementary lecture notes will be posted there for student use. Note that reading the supplementary lecture notes only is NOT ENOUGH to pass this course! The single best predictor of student performance is attendance at lectures.

Grading Scheme:
Problem Assignments: ........................................ = 20 %
Literature Review(s): ...................................... = 20 %
Term project: ................................................. = 30 %
Final Exam: .................................................... = 30 %

Note that only portions of the assignment(s) may be used for purposes of evaluation.
**Important Notes:**

1) **Students are expected to attend all lectures.** If a student is absent from a lecture, it is up to the student to obtain missed lecture material from colleagues in the course.

2) **Students who miss a deadline** due to illness must provide a valid medical certificate to the instructor not later than 48Hrs after returning to campus. The certificate must clearly state the name of the doctor with contact information, the degree of incapacitation, the time & date of onset, the time & date that you were seen, and the expected recovery date. Once the certificate has been verified, alternate arrangements will be negotiated with the instructor.

Students who miss a deadline without a very good reason may be denied the right to submit their work or may be penalized (e.g. 10% per day).

3) **Academic Accommodation.** You may need special arrangements to meet your academic obligations during the term because of disability, pregnancy or religious obligations. Please review the course outline promptly and write to me with any requests for academic accommodation during the first two weeks of class, or as soon as possible after the need for accommodation is known to exist.

Students with disabilities requiring academic accommodations in this course must register with the Paul Menton Centre for Students with Disabilities (PMC) for a formal evaluation of disability-related needs. Documented disabilities could include but not limited to mobility/physical impairments, specific Learning Disabilities (LD), psychiatric/psychological disabilities, sensory disabilities, Attention Deficit Hyperactivity Disorder (ADHD), and chronic medical conditions. Registered PMC students are required to contact the PMC, 613-520-6608, every term to ensure that your Instructor receives your Letter of Accommodation, no later than two weeks before the first assignment is due or the first in-class test/midterm requiring accommodations. If you only require accommodations for your formally scheduled exam(s) in this course, please submit your request for accommodations to PMC by the last official day to withdraw from classes in each term.

4) **Plagiarism.** Plagiarism and cheating at the graduate level are viewed as being particularly serious and the sanctions imposed are accordingly severe. Students are expected to familiarize themselves with and follow the Carleton University Student Academic Integrity Policy (See https://carleton.ca/registrar/academic-integrity/). The Policy is strictly enforced and is binding on all students. Plagiarism and cheating – presenting another’s ideas, arguments, words or images as your own, using unauthorized material, misrepresentation, fabricating or misrepresenting research data, unauthorized cooperation or collaboration or completing work for another student – weaken the quality of the graduate degree. Academic dishonesty in any form will not be tolerated. Students who infringe the Policy may be subject to one of several penalties including: expulsion; suspension from all studies at Carleton; suspension from full-time studies; and/or a reprimand; a refusal of permission to continue or to register in a specific degree program; academic probation; or a grade of Failure in the course.

5) **Health and Safety.** Every student should have a copy of our Health and Safety Manual. An electronic version of the manual can be found at: [http://www.sce.carleton.ca/courses/health-and-safety.pdf](http://www.sce.carleton.ca/courses/health-and-safety.pdf)

**Week-by-Week Outline (subject to change):**

**Weeks 1-2:** Introduction to Pattern Classification. Data pre-processing, analysis, outlier detection, and transformations. Experiment design (feature selection & dimensionality, selecting classifier structure, test protocols, cross-validation, data partitioning, etc.). Introduction to hypothesis testing. Avoiding fundamental errors of testing on the training set and training on the test set.

**Week 3:** Reporting results. How to accurately and honestly report classification system performance. True error vs. apparent error. Confidence intervals, statistical tests to compare methods, receiver operator characteristic curves, sensitivity, specificity, confusion matrices, P-values. Class imbalance. Critical assessment of reported results in the literature.

Note that by introducing experiment design and reporting of results early in the course, applications discussed during all subsequent topics will be evaluated using these fundamental principles. The remainder of the course
will survey a number of approaches to pattern classification. The depth of coverage will vary and will depend on time available. Relevant applications of pattern classification techniques from the literature will be discussed.


**Week 5:** Maximum likelihood and Bayesian parameter estimation. Non-parametric techniques such as Parzen windows, probabilistic neural networks, and K-nearest neighbour estimators and classifiers.

**Week 6:** Decision trees and decision forests. Training, pruning, splitting, and stopping criteria.

**Week 7:** Linear and nonlinear discriminant analysis. Linear discriminant functions and decision surfaces. Perceptron criterion, relaxation procedures, and MSE procedures. Generalized linear discriminant functions and support vector machines (briefly).

**Week 8:** Neural networks: network structure, feedforward operation and classification, backpropagation training.

**Week 9:** Markov chains, hidden Markov models, and expectation maximization.

**Week 10:** Meta-learner and re-sampling approaches including bagging and boosting. Combination of multiple experts: voting strategies and cascaded classifiers. Learning with queries.

**Week 11:** Unsupervised clustering (hierarchical, K-means, SOMs). Mixture densities, criterion functions for clustering, the number of cluster problem, and cluster validation.

**Week 12/13:** Student project presentations and competition. Review.