Programming for Cognitive Science and Artificial Intelligence

COGS 4410, COGS 6410, CSCI 4966

Course Objective

Research in Cognitive Science and Artificial Intelligence (AI) is driven by data. Researchers in these fields collect, manipulate, model and analyze data generated by real world processes. Since the amount of data available has grown exponentially, the ability to automate these tasks through computer programs is essential. Specifically, we need probabilistic and statistical computing, to learn from the data and to handle the uncertainty inherent in the data. The objectives of this course are for the student to acquire the basics of statistical and machine learning and proficiency in probabilistic and statistical programming. They will be able to transform, visualize and model data acquired from real world datasets.

Course Format

The course format is a combination of lecture, class discussions and in-class exercises. The homework will consist of developing programs that learn from real world datasets. The Python programming language, data processing and machine learning libraries will be used.

Course Learning Outcomes for COGS-4410 and CSCI-4966

Students who successfully complete this course will be able to:

- Demonstrate the ability to preprocess and visualize data for modeling.
- Demonstrate the ability to generate regression and classification models.
- Demonstrate the ability to generate tree-based models.
- Demonstrate the ability to generate neural network models including some deep learning models.
- Demonstrate the ability to perform model selection and validation

Course Learning Outcomes for COGS-6410

Students who successfully complete this course will be able to:

- Demonstrate the ability to preprocess and visualize data for modeling.
- Demonstrate the ability to generate regression and classification models.
- Demonstrate the ability to generate tree-based models.
- Demonstrate the ability to generate neural network models including some deep learning models.
- Demonstrate the ability to perform model selection and validation.
- Demonstrate knowledge of statistical and machine learning theory
- Demonstrate knowledge of the mathematics underlying statistical and machine learning

Instructor

Michael Schoelles

Office: https://rensselaer.webex.com/meet/schoem

Email: schoem@rpi.edu

Office Hours: Virtual Meeting, email for time

Teaching Assistant

None

Meeting Time and Place

Tuesday/Friday, 10:10 - 12, Online

Learning Management System

https://lms.rpi.edu

Materials

Python 3.6 or higher (not Python 2.x) will be used in the course

Basic free Anaconda install with Python:

https://www.anaconda.com/download/ (includes most of the required libraries)

Course Assessment Measures

COGS 4410	% of Grade	COGS 6410	% of Grade
CSCI 4966			
Homework	80	Homework	80
Class Participation	20	Class Participation	20

Homework1: Regression Methods 20%

Homework2: Classification Methods 20%

Homework3: Ensemble Methods 20%

Homework4: Deep Learning Methods 20%

Attendance Policy

I strongly encourage and expect you to attend class.

Homework Policy

All work is submitted via the LMS system. Students must work alone on homework submissions. There will be a 10% penalty for late homework and exercises. Homework and exercises will not be accepted after 1 week from due date.

Exam Policy

All exams require some coding so the exam is done on your computer and the use of the internet is allowed for programming references. Use of class notes is permitted.

Academic Integrity

Student-teacher relationships are built on trust. For example, students must trust that teachers have made appropriate decisions about the structure and content of the courses they teach, and teachers must trust that the assignments that students turn in are their own. Acts that violate this trust undermine the educational process. The Rensselaer Handbook of Student Rights and Responsibilities and The Graduate Student Supplement define various forms of Academic Dishonesty and you should make yourself familiar with these. In this class, all assignments that are turned in for a grade must represent the student's own work. In cases where help was received, or teamwork was allowed, a notation on the assignment should indicate your collaboration.

Submission of any assignment that is in violation of this policy will result in the following penalties. For the first violation a grade of 0 will be given for that assignment. The second violation results in failure of the course and the student will be reported to the Dean of Students or the Dean of Graduate Education. If you have any questions concerning this policy before submitting an assignment, please ask for clarification.

Students with Disabilities:

Rensselaer Polytechnic Institute strives to make all learning experiences as accessible as possible. If you anticipate or experience academic barriers based on a disability, please let me know immediately so that we can discuss your options. To establish reasonable accommodations, please register with The Office of Disability Services for Students. After registration, make arrangements with me as soon as possible to discuss your accommodations so that they may be implemented in a timely fashion. DSS contact information: dss@rpi.edu; 518-276-819; 4226 Academy Hall.

Schedule (Tentative and Flexible)

Days	Topic	Homework
Sept 1/4	Introduction	
	Exploratory Data Analysis &	
	Data Preprocessing	
Sept 11	Exploratory Data Analysis&	
	Data Preprocessing	
	Linear Regression	
Sept 15/18	Linear Regression	
	Probability/Statistics	
Sept 22/25	Logistic Regression	
	Neural Network Logistic Regression	
Sept 29/Oct 2	Linear Algebra	Homework1 (due Sept 29)
	Support Vector Machines	
Oct 6/9	Support Vector Machines	
	Class Canceled (10/9)	
Oct 13/16	Generative Models	Homework 2 (due Oct 16)
	Naïve Bayes	
	Linear Discriminant Analysis	
	Model Selection	
	Bias-Variance Tradeoff	
	Cross -Validation	
	Regularization	
Oct 20/23	Model Selection	
	Feature Selection	
	Metrics	
	Decision Trees	
	Bagging	
	Random Forest	
Oct 27/30	Boosting	
	ADABoost	
	Gradient Boosting	
	XGBoost	
	Catboost	
Nov 3/6	Clustering	
	Dimension Reduction	
Nov 10/13	Deep Learning	
	Tensors	
	Autograd	
	Optimizers	
	CNN	
Nov 17/20	Deep Learning	Homework 3 (due Nov 17)
	ResNet	
	RNN,LSTM,GRU	
Nov 24/27	No Class	
Dec 1/4	Deep Learning	
	Autoencoders	
	GANS	
Dec 8/11	Deep Learning	Homework 4 (due Dec 12)
	Reenforcement Learning	
	NLP	