Course Syllabus: CPCSI 791TS
Machine Learning and Time Series – Spring 2013

Instructor: Benjamin M. Marlin
Email: marlin@cs.umass.edu
Office: CS234
Office Hours: TBA

First Class: January 23, 2013
Time: MW 2:05-3:20
Location: CS 142

Course Description: This seminar will focus on models and algorithms for supervised and unsupervised machine learning with time series. Topics will include discrete and continuous time models from machine learning, statistics and econometrics. We will investigate a variety of time series problems including prediction, detection, clustering, and similarity search. Coursework for the one credit option will include paper presentations and quizzes. Students in the three credit option will also complete a course project. Prerequisites include a graduate-level course in machine learning or graphical models, i.e., CMPSCI 689 or CMPSCI 691GM.

Textbooks and References: This course will use material from several sources. For statistical background on time series modeling, our primary reference will be *Time Series Analysis* by Cryer and Chan. This text is available at no cost to UMass students through Springer Link here. As a secondary statistical text, we may also use *Time Series Analysis and Its Applications* by Shumway and Stoffer. This text is available at no cost to UMass students through Springer Link here. Students may also find it useful to have references for graphical models and machine learning. Recommended texts are *Probabilistic Graphical Models* by Koller and Freidmann (the course text for 691GM), and *Machine Learning: a Probabilistic Perspective* by Murphy (the course text for 689).

Course Website: The course website is hosted on UMass's Moodle course management portal https://moodle.umass.edu/. You log into the portal using your OIT user name and password. The course website will host pointers to readings and demos, announcements, and discussion forums.

Announcements: Official announcements for the course will go out through the Moodle portal as email and will be automatically logged as news items. Email will be sent to your *official UMass email address* (@student.umass.edu, @cns.umass.edu, etc...).

Grading Plan: The coursework for the one credit option will consist of leading and participating in discussions about readings, preparing and presenting demos, and completing in-class quizzes. The coursework for the three credit option will also include a substantial course project applying machine learning and/or statistical methods to model or analyze time series data. The project components will include a written proposal (2 pages), a proposal presentation (10 minutes plus questions), a final report (8-10 pages) and a final presentation (20 minutes questions). The grading schemes are given below:

Activity	Section A (3 credits)	Section B (1 credit)
Reading and Demo Presentations	10%	30%
Discussion Participation	10%	20%
Quizzes	15%	50%
Project Proposal Presentation	15%	
Project Final Report	50%	

Course Policies:

- **Course Participation:** This is a discussion-based seminar. You must be in class to participate in and lead discussions. If you are going to miss class on a day where you are assigned to present a reading or demo, let the instructor know **in advance** so the presentation schedule can be modified. If you unexpectedly miss a class where you are scheduled to present due to illness or other exceptional circumstance, you must provide a doctor's note or other supporting documentation at your earliest convenience. You will need to make up for missed presentations by presenting new material on an alternate day.
- **Reading Presentations:** You may give your reading presentations using slides or the whiteboard. You may use any source materials you like in addition to the assigned reading to prepare for your presentation. Consulting additional texts or web pages to gain a better understanding of the assigned material is encouraged. However, you must cite your sources. You are also encouraged to attend office hours before your presentation if you have any questions about the assigned material.
- **Demo Presentations:** You may create your demo in any programming language you like. However, Python, Matlab/Octave and R are strongly recommended since they have very good plotting packages as well as good statistical packages. Consulting additional texts or web pages to gain a better understanding of the material needed to implement a demo is encouraged. However, you must cite your sources.
- **Late Project Components:** Project components must be submitted by the date they are due unless an extension is arranged with the instructor **in advance** of the due date. Extensions may be granted due to conflicts with the student's primary research (ex: conflicting conference publication deadlines).
- **Academic Honesty:** Students must abide by the University's academic honesty policy at all times. Academic dishonesty includes but is not limited to: cheating, fabrication, plagiarism, and facilitating dishonesty.

Important Dates – Spring 2013:

Jan 22, 2013	First day of classes
Feb 4, 2013	Add/Drop date
Feb 18, 2013	Holiday – President's Day
Feb 19, 2013	Monday class schedule
Mar 4, 2013	Last day for grad students to drop with DR
Mar 7, 2013	Last day for undergrads to drop with W
Mar 16, 2013	Spring Break Starts
Mar 25, 2013	Classes Resume
Apr 15, 2013	Holiday – Patriot's Day
Apr 17, 2013	Monday Schedule
May 1, 2013	Last Day of classes

Approximate Schedule: (Subject to change over the semester)

Meeting	Date	Topics
Meeting 1	We Jan 23, 2013	Course logistics and overview. Probability Review.
Meeting 2	Mo Jan 28, 2013	White Noise, Random Walks, Moving Averages and Stationarity
Meeting 3	We Jan 30, 2013	Estimating Trends: Constant Mean and Regression Methods
Meeting 4	Mo Feb 4, 2013	General Linear Processes and Moving Average Processes
Meeting 5	We Feb 6, 2013	Autoregressive Processes and Mixed Models
Meeting 6	Mo Feb 11, 2013	Autoregressive Integrated Moving Average (ARIMA) Models
Meeting 7	We Feb 13, 2013	ARIMA Models and Maximum Likelihood Estimation
	Mo Feb 18, 2013	No class – President's Day
Meeting 8	Tu Feb 19, 2013	ARIMA Models and Bayesian Inference
Meeting 9	We Feb 20, 2013	ARIMA Models and Forecasting
Meeting 10	Mo Feb 25, 2013	Autoregressive Conditional Heteroskedasticity (ARCH) Models
Meeting 11	We Feb 26, 2013	Time Series Regression Models
Meeting 12	Mo Mar 4, 2013	Gaussian Processes and Gaussian Markov Random Fields
Meeting 13	We Mar 6, 2013	Model Comparisons Discussion
Meeting 14	Mo Mar 11, 2013	Project Proposal Presentations
Meeting 15	We Mar 13, 2013	Project Proposal Presentations
	Mo Mar 18, 2013	No class – Spring Break
	We Mar 20, 2013	No class – Spring break
Meeting 16	Mo Mar 25, 2013	Dynamic Bayesian Networks
Meeting 17	We Mar 27, 2013	Continuous Time Bayesian Networks
Meeting 18	Mo Apr 1, 2013	Time Series Similarity: L2 Distance and Dynamic Time Warping
Meeting 19	We Apr 3, 2013	Time Series Similarity: Symbolic Approximation (SAX)
Meeting 20	Mo Apr 8, 2013	Time Series Classification
Meeting 21	We Apr 10, 2013	Time Series Detection and Labeling
Meeting 22	Mo Apr 15, 2013	Time Series Clustering
Meeting 23	We Apr 17, 2013	Time Series Indexing and Retrieval
Meeting 24	Mo Apr 22, 2013	Final Project Presentations
Meeting 25	We Apr 24, 2013	Final Project Presentations
Meeting 26	Mo Apr 29, 2013	Course wrap up