CIS 732: Machine Learning and Pattern Recognition
CIS 830: Topics in Artificial Intelligence
Spring 2016

Hours: 3 hours (additional 3-hour proseminar in data mining available as CIS 798 or CIS 890)
Prerequisite: CIS 300 and 501 or equivalent coursework in data structures and algorithms; CIS 301 (set theory/logic), Math 510 (discrete math), Stat 410 (intro probability) recommended
Textbook: Murphy, K. J. (2012). *Machine Learning: A Probabilistic Perspective.* Cambridge, MA: MIT Press.
Time and Venue: Mon, Wed, Fri 13:00 – 14:20, Room 1063 Rathbone Hall
Instructor: William H. Hsu, Department of Computing and Information Sciences
Office: 2164 DUE Phone (Google Voice: office/home/cell): +1 785 236 8247
E-mail: bhsu@ksu.edu URL: http://www.cis.ksu.edu/~bhsu TA: Pavel Janovsky
Instructional e-mail alias (to reach instructor & TA): CIS732TA-L@listserv.ksu.edu
Office hours: 11:30-12:30 Mon, 10-11 Tue, 12:00-13:00 Wed; 10:30-11:30 Fri; by appointment
Class web page: http://bit.ly/kstate-mlclass-2016 (Canvas), http://bit.ly/kstate-mlclass (public)

Course Description

This is an introductory course in machine learning for development of intelligent knowledge based systems. The first half of the course will focus on basic taxonomies and theories of learning, algorithms for concept learning, statistical learning, knowledge representation, pattern recognition, and reasoning under uncertainty. The second half of the course will survey fundamental topics in combining multiple models, learning for plan generation, decision support, knowledge discovery and data mining, control and optimization, and learning to reason.

Course Requirements

Exams (35%): in-class midterm exam (15%), take-home final (20%)
Homework (32%): 8 out of 10 programming and written assignments (4% each: 3 written, 3 programming, 4 mixed)
Project (20%): term programming project and report for all students
Paper Reviews (10%): 10 of 13 weekly or semi-weekly paper reviews (1% each)
Class Participation (3%): class and online discussions, asking and answering questions

Computer language(s): C/C++/C#, Java, or student choice (upon instructor approval)

Selected reading (on reserve in K-State CIS Library)

- Alpaydin. E. (2014). *Introduction to Machine Learning, 3e.* Cambridge, MA: The MIT Press.
- Han, J., Kamber, M., & Pei, J. (2012). *Data Mining: Concepts and Techniques, 3e.* San Francisco, CA: Morgan Kauffman.
- Haykin, S. (2009). *Neural Networks and Learning Machines, 3rd edition*. Englewood Cliffs, NJ: Prentice-Hall.
- Koza, J. R., Keane, M. A., Streeter, M. J., Mydlowec, W., Yu, J., & Lanza, G. (2005). *Genetic Programming IV: Routine Human-Competitive Machine Intelligence*. New York, NY: Springer.
- Mitchell, T. M. (1997). *Machine Learning.* New York, NY: McGraw-Hill.
- Bishop, C. M. (1995). *Neural Networks for Pattern Recognition*. London, UK: Oxford University Press.
- Goldberg, D. E. (1989). *Genetic Algorithms in Search, Optimization, and Machine Learning*. Reading, MA: Addison-Wesley.
- Koza, J. (1992). *Genetic Programming: On The Programming of Computers by Means of Natural Selection*. Cambridge, MA: The MIT Press.

Course Calendar

Lecture	Date	Topic	Primary Source: Murphy
0	Wed 20 Jan 2016	Administrative; overview of learning	Syllabus; §1.1 – 1.2
1	Fri 22 Jan 2016	Supervised/unsupervised learning intro	§1.2 – 1.3
2	Mon 25 Jan 2016	Unsupervised learning: basics; LDA (830)	§1.3 – 1.4; §12.1 – 12.2 (830)
3	Wed 27 Jan 2016	Probability concepts 1 of 5: review of basics	Chapter 2 (esp. §2.1 – 2.7)
4	Fri 29 Jan 2016	Information theory basics	Handout 1; §2.8
5	Mon 01 Feb 2016	Probability concepts 2 of 5: intro to Bayes	§3.1 – 3.4
6	Wed 03 Feb 2016	Naïve Bayes & applications; clustering	Handout 2; §3.5; Chapter 25
7	Fri 05 Feb 2016	Probability concepts 3 of 5: Bayesian stats	Chapter 5 (esp. §5.1 – 5.4)
8	Mon 08 Feb 2016	Hierarchical models & loss functions	Handout 3; §5.5 – 5.7
9	Wed 10 Feb 2016	Probability concepts 4 of 5: frequentist stats	Chapter 6
10	Fri 12 Feb 2016	Linear regression: intro	§7.1 – 7.3
11	Mon 15 Feb 2016	Linear regression: ridge, robust, Bayesian	§7.4 – 7.6
12	Wed 17 Feb 2016	Logistic regression: intro	§8.1 – 8.3
13	Fri 19 Feb 2016	Logistic regression: generative, discriminative	Handout 4; §8.4 – 8.6
14	Mon 22 Feb 2016	Neural networks: intro (perceptrons, Winnow)	§8.6; Mitchell Chapter 4
15	Wed 24 Feb 2016	Neural networks: feedforward/MLP, backprop	§16.5; Mitchell Chapter 4
16	Fri 26 Feb 2016	EM: robust & probit regression	Handout 5; §9.4; §11.1 – 11.4
17	Mon 29 Feb 2016	Probability concepts 5 of 5: Gaussian models	Chapter 4 (esp. §4.1 – 4.4)
18	Wed 02 Mar 2016	Generative classifiers; discriminant analysis	Handout 6; §3.1 – 3.2; §8.6
19	**Fri 04 Mar 2016**	**Mixture models; exam review**	**Chapter 11 (esp. §11.1 – 11.4)**
20	Mon 07 Mar 2016	Sparse linear models	Chapter 13 (esp. §13.1 – 13.2)
21	Wed 09 Mar 2016	L1 regularization	§13.3 – 13.5
	Fri 11 Mar 2016	**Online/take-home midterm (732 & 830)**	**1 – 8, §9.4, 11 – 13, §16.5, 25**
22	**Mon 21 Mar 2016**	**Variable selection; exam solution review**	**Handout 7; §13.2**
23	Wed 23 Mar 2016	Sparse kernel machines	Chapter 14 (esp. §14.1 – 14.5)
24	Fri 25 Mar 2016	Sparse kernel machines, Gaussians	§14.6 – 14.7; §15.1 – 15.2
25	Mon 28 Mar 2016	Gaussian processes	§15.3 – 15.6
26	Wed 30 Mar 2016	Bayesian inference	Handout 8; Chapter 10 (830)
27	Fri 01 Apr 2016	Bayesian stats intro; Bayesian networks (830)	Chapter 5; Chapter 10 (830)
28	Mon 04 Apr 2016	More Bayesian stats; more BNs (830)	Chapter 5; Chapter 10 (830)
29	Wed 06 Apr 2016	Empirical, hierarchical, variational Bayes	§5.5 – 5.6; §21.1 – 21.5
30	Fri 08 Apr 2016	More variational Bayes	Handout 9; §21.5
31	Mon 11 Apr 2016	Graphical models (GM)	Chapter 10 (everyone)
32	Wed 13 Apr 2016	Hidden Markov models & Kalman filters	Chapter 17 (esp. §17.1 – 17.4)
33	Fri 15 Apr 2016	Forward-backward algorithm	§17.4 – 17.5; §23.1 – 23.4
34	Mon 18 Apr 2016	Monte Carlo; Gibbs, simulated annealing	Chapter 23; §24.1 – 24.3
35	Wed 20 Apr 2016	Deep learning intro; GM inference (830)	Handout 10; Chapter 20 (830)
36	Fri 22 Apr 2016	Deep learning: autoencoders	Chapter 28
37	Mon 25 Apr 2016	Applications in data science; transfer learning	Handout 11
38	Wed 27 Apr 2016	Semi-supervised, active, transfer learning	Handout 11; Chapter 9
39	**Fri 29 Apr 2016**	**GEC overview; GM structure (830); final**	**Handout 12; Chapter 26 (830)**
40	Mon 02 May 2016	GEC: genetic algorithms	Handout 12
41	Wed 04 May 2016	GEC: genetic programming	Handout 12
42	**Fri 06 May 2016**	**Term project data blitzes**	**1 – 17, 20 – 21, 23 – 25, 26, 28**
		732 Take-Home Final due	**1 – 17, 20 – 21, 23 – 25, 26, 28**

Green-shaded entries denote the due date of a paper review.
Lightly-shaded entries denote the due date of a written problem set – after Lectures 4, 20, and 34.
Intermediate-shaded entries denote the due date of a mixed homework – 11, 25, 28, 34.
Heavily-shaded entries denote the due date of a machine problem (programming assignment) – 8, 17, 31.

Interim project interviews will be held between the midterm and spring break.
The blue-shaded date is the due date of the draft project report and demo, with interviews and presentations to be held the last two days of class.
Green, **blue** and **red** letters denote **exam review**, **exam**, and **exam solution review** dates.
Underlined topics are covered only in 830.
Orange letters denote handouts distributed via K-State Online (KSOL)'s Canvas, and via the public mirror.