QMB 7933 - Fall 2017, Mod 1

PhD Seminar in IS
Learning Theory & Methodologies

Haldun Aytug
STZ 351
392-2468
aytugh@ufl.edu
Office hours: TBA

1. Objectives

The main objective of this course is to achieve a deeper understanding of the research problems associated with machine learning (data mining) theory, methodologies and some novel applications. The course will cover basics of regression (classification as regression), classification, clustering and methodologies involved in estimation of regression, classification and clustering functions. The course will draw upon knowledge of probability theory, statistics, linear algebra, and optimization theory.

We will cover material from multiple textbooks and papers. Some of the notation and content is from Cristianini & Shawe-Taylor (2000) and some from Vapnik (1998), as well as Hastie et al. (2009) and Shalizi (2013) (see http://www.stat.cmu.edu/~cshalizi/ADAfaEPoV/). We will also discuss a small portion of Goodfellow et al. (2016).

2. Course Procedures

The course is a mixture of lectures (by me) and presentations (by you).

This is a PhD level course which requires substantial work outside the classroom. From time to time, I will assign questions to be solved (some might require analyzing data) and handed in next time we meet (no exceptions). I plan to have at least one quiz. The exam is cumulative.

Each student will be asked to present a summary of findings and open research questions on one of the topics suggested in Section 7.

To complete the course, the student has to take the final exam and quizzes, complete assignments, and write a satisfactory summary report on his/her selected topic.

3. Class Meetings

We will meet on Fridays, starting at 1pm. for 3-4 hours in room STZ 103. No class on Sept 22nd. Instead we will meet September 11th at 1 pm in STZ 200. We will also reschedule the homecoming lecture of October 6th (date, time and location TBD).

4. Grading

Your final grade will depend on your performance on

- Final Exam - 30%
- Proposal and presentation - 30%
- Assignments - 25%
- Quizzes - 15%

5. Course Schedule

1. Week 1 - The regression function, bias variance trade-off. - (Shalizi, 2013, Ch. 1), (Hastie et al., 2009, Ch. 2).
2. Week 2 - Nearest neighbor regression, kernel regression, linear regression smoothers. - (Shalizi, 2013, Ch 1,2), (Hastie et al., 2009, Ch. 2,3).
3. Week 2 - Regression with subset selection (ridge regression, the LASSO). (Hastie et al., 2009, Ch. 3).
4. Week 3 - Classification. - (Hastie et al., 2009, Ch. 4), (Shalizi, 2013, Ch 12).
5. Week 3 - Support Vector Machines, Optimization Theory - (Cristianini & Shawe-Taylor, 2000, Ch. 5), (Bazaraa et al., 1993)
6. Week 4 - Regularization, splines - (Hastie et al., 2009, Ch. 5), (Shalizi, 2013, Ch 8).
7. Week 4- Kernels - (Hastie et al., 2009, Ch. 5), (Shalizi, 2013, Ch 4)
8. Week 4 - Kernel Machines - (Cristianini & Shawe-Taylor, 2000, Ch. 3), (Mangasarian, 2000), (Shawe-Taylor & Cristianini, 2004, Ch. 3,9)
9. Week 5 - Deep learning - (Goodfellow et al., 2016)
10. Week 6 - Model Assessment - (Hastie et al., 2009, Ch. 7), (Shalizi, 2013, Ch 3)
11. Week 6 - Introduction to Bayesian approaches - TBD
12. Week 6 - Clustering - TBD
13. Week 7 - Intro to Statistical Learning Theory - (Vapnik, 1998, Ch. 1, 4),(Cristianini & Shawe-Taylor, 2000, Ch. 4), (Shawe-Taylor et al., 1998), (Scholkopf et al., 1999, Ch 4)
14. Week 8 - Presentations (Research Topic)
15. Exam

6. List of Papers to read and present

Please pick one of the papers to present. This is in addition to your group report and presentation discussed in Section (7). The common thread across these articles is that most use or propose a methodology (for prediction) relevant to what we are discussing here and most have business related applications.

1. Saar-Tsechansky & Provost (2007)
2. Ban et al. (2016)
3. Harrison et al. (2012)
4. Atahan & Sarkar (2011)
5. Zheng & Padmanabhan (2006)
6. Cui et al. (2006)
7. Das et al. (1999)
8. Lee et al. (2016)
9. Chen et al. (2017)
10. Ching et al. (2013)
11. Zhao et al. (2013)
12. Dzyabura & Hauser (2011)
13. Meyer et al. (2014)
14. Yang et al. (2013)

7. List of Research Topics

I have listed below several potential research topics. I will ask you to choose a topic in teams of two, read the literature, and write a proposal outlining what has been done in the field, what you think are open research questions.

- Support Vector Regression
- Classification on time series and/or sequential data, Hidden Markov Models
- Utility based learning
- Strategic (Adversarial) Learning
- Social Network Analysis

- Regularization theory
- Generalized Linear Models
- Hierarchical Linear Models
- Feature Selection
- PCA, Generalized PCA
- Deep Learning
- MCMC
- Bayesian Learning
- BIG DATA
- Text Mining, Topic Modeling (LDA)
- ...

References

Atahan, P. & Sarkar, S. (2011). Accelerated learning of user profiles. *Management Science*, **57**(2), 215–239.

Ban, G.-Y., Karoui, N. E., & Lim, A. E. B. (2016). Machine learning and portfolio optimization. *Management Science*, (0), null.

Bazaraa, M., Sherali, H., & Shetty, C. (1993). *Nonlienar Programming, Theory and Algorithms*. John Wiley and Sons, Inc, New York.

Chen, Y., Iyengar, R., & Iyengar, G. (2017). Modeling multimodal continuous heterogeneity in conjoint analysisa sparse learning approach. *Marketing Science*, **36**(1), 140–156.

Ching, A. T., Erdem, T., & Keane, M. P. (2013). Invited paperlearning models: An assessment of progress, challenges, and new developments. *Marketing Science*, **32**(6), 913–938.

Cristianini, N. & Shawe-Taylor, J. (2000). *An Introduction to Support Vector Machines and other kernel-based learning methods*. Cambridge University Press, United Kingdom.

Cui, G., Wong, M. L., & Lui, H.-K. (2006). Machine learning for direct marketing response models: Bayesian networks with evolutionary programming. *Management Science*, **52**(4), 597–612.

Das, T. K., Gosavi, A., Mahadevan, S., & Marchalleck, N. (1999). Solving semi-markov decision problems using average reward reinforcement learning. *Management Science*, **45**(4), 560–574.

Dzyabura, D. & Hauser, J. R. (2011). Active machine learning for consideration heuristics. *Marketing Science*, **30**(5), 801–819.

Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep Learning*. MIT Press, Cambridge, MA.

Harrison, J. M., Keskin, N. B., & Zeevi, A. (2012). Bayesian dynamic pricing policies: Learning and earning under a binary prior distribution. *Management Science*, **58**(3), 570–586.

Hastie, T., Tibshirani, R., & Friedman, J. (2009). *The Elements of Statistical Learning*. Springer.

Lee, E. K., Nakaya, H. I., Yuan, F., Querec, T. D., Burel, G., Pietz, F. H., Benecke, B. A., & Pulendran, B. (2016). Machine learning for predicting vaccine immunogenicity. *Interfaces*, **46**(5), 368–390.

Mangasarian, O. L. (2000). *Advances in Large Margin Classifiers*, chapter Generalized Support Vector Machines. MIT Press.

Meyer, G., Adomavicius, G., Johnson, P. E., Elidrisi, M., Rush, W. A., Sperl-Hillen, J. M., & O'Connor, P. J. (2014). A machine learning approach to improving dynamic decision making. *Information Systems Research*, **25**(2), 239–263.

Saar-Tsechansky, M. & Provost, F. (2007). Decision-centric active learning of binary-outcome models. *Information Systems Research*, **18**(1), 4–22.

Scholkopf, B., Burges, C., & Smola, J. (1999). *Advances in Kernel Methods. Support Vector Learning*. MIT Press; Cambridge, MA.

Shalizi, C. R. (2013). *Advanced Data Analysis from an Elementary Point of View*.

Shawe-Taylor, J. & Cristianini, N. (2004). *Kernel Methods for Pattern Analysis*. Cambridge University Press, United Kingdom.

Shawe-Taylor, J., Bartlett, P., Williamson, R., & Anthony, M. (1998). Structural risk minimization over data-dependent hierarchies. *IEEE Transactions on Information Theory*, **44**(5), 1926–1940.

Vapnik, V. (1998). *Statistical Learning Theory*. John Wiley and Sons, New York.

Yang, Y. C., Liu, H., & Cai, Y. (2013). Discovery of online shopping patterns across websites. *INFORMS Journal on Computing*, **25**(1), 161–176.

Zhao, Y., Yang, S., Narayan, V., & Zhao, Y. (2013). Modeling consumer learning from online product reviews. *Marketing Science*, **32**(1), 153–169.

Zheng, Z. & Padmanabhan, B. (2006). Selectively acquiring customer information: A new data acquisition problem and an active learning-based solution. *Management Science*, **52**(5), 697–712.