CS671-Machine Learning
FALL 2017

Prof. Dan A. Simovici

Office hours: MW 5:30-6:30 pm

This page is posted on www.cs.umb.edu/~dsim; on the same site you will find copies of the slides I am using in class, homeworks, and handouts relevant to the course. You should visit it often!

Machine Learning is a mathematical discipline that is the foundation for Data Mining and Data Analysis. I expect that you had previous courses in linear algebra and probabilities, but everything possible will be done to make the course self-contained. The main sources for this course are:

Foundations of Machine Learning by Mehryar Mohri, Afshin Rostamizadeh and Ameet Talwalkar, MIT Press, 2012;

Understanding Machine Learning by S. Shalev-Shwartz and S. Ben-David, Cambridge University Press, 2016;

Y. S. Abu-Mostafa, M. Magdon-Ismail, and H.T. Lin, Learning from Data, AMLbook.com, 2012;

M. J. Kearns and U. V. Vazirani, An Introduction to Computational Learning Theory, MIT Press, 1997;

K. P. Murphy: Machine Learning - A Probabilistic Perspective, MIT Press, 2012;

M. Anthony and N. Biggs: Computational Learning Theory, Cambridge University Press, 1997.

D. Simovici and C. Djeraba: Mathematical Tools for Data Mining, 2nd edition, Springer 2014.

In addition, I will use several classical probabilities texts. Your primary source for this course should be the slides that I will present in class and place on the web site.

We shall cover the following topics:

1. Probably Approximately Correct Learning
2. The Vapnik-Chervonenkis Dimension
3. Support Vector Machines
4. Boosting
5. Regression
6. Data Provenance

Homework should be entirely the product of your work; you may discuss it with colleagues and I encourage you to come to talk to me if you have difficulties. Also, learn LaTeX and use it to write your homework. Homework and class participation determine your grade.

Machine Learning is a wonderful subject and I hope that you will enjoy it!