

Blogroll**Social**

index

Permalink: [2019-01-23 00:00:00-05:00](#) by [Mikael Vejdemo-Johansson](#) in [misc](#)

Welcome to MTH 513

We will be covering Machine Learning, using a competition based approach to the topic. The class will be divided into teams and participate in competitions on [Kaggle](#).

The course meets Mondays and Wednesdays, 10.10-12.05, in 1S-103.

Your instructor is Prof. Mikael Vejdemo-Johansson. On this website you will be able to find lecture notes for the lectures given during the course. Most of the course will be committed to lab time working on the competitions, with lectures an occasional interspersed activity.

We will use [An Introduction to Statistical Learning](#) as a reference book. It is available for free as an ebook, and through the library you can find it on SpringerLink, where a printed copy can be ordered for \$25.

Contact

Mikael.VejdemoJohansson@csi.cuny.edu
1S-208
Office hours Monday, Wednesday, 13.00 - 14.30.

Competitions

- [Titanic](#) - MVJ got 0.80382
- [Tabular Playground Series - Feb 2022](#) - MVJ got 0.82119
- [NYC Taxi Trip Duration](#) - MVJ got 0.76346
- [Petals to the Metal](#) - MVJ got 0.22793

Example solutions

[michiexile kernels](#)

Course Blog

The course is primarily graded on writing blog posts describing, in detail, the process and the solution for one of the team's competition submissions.

The blog is [here on Medium](#).

A blog post should include a narrative exposition of the entire process for its competition. In particular it should include

- What models were tried? What did they score? Why were models or model components discarded?
- What auxiliary information (exploratory plots etc) were used to decide?
- What data transformations were tried? What determined which ones to keep and which ones to discard?
- What was the final pipeline? What did it score?

To create your submission, first make sure you have a Medium account. Next, send your username to the professor so that you can be included in the blog writing crew. Finally write up your blog post and when you are done, use the [...] menu at the bottom to submit it to the course blog publication.

The blog post is due at midnight following the final exam. Feedback will be given if a draft is shown at least a week earlier.

Final Exam

The final exam is May 18, 10.10 - 12.05. For the final, any handwritten notes you bring are allowed, but no other documentation.

Lecture slides

These slides are based on the slides written for the 2019 course. Topics appear in a different order this time around, which is occasionally reflected in the slides.

1	2022-01-31 Welcome, basic concepts of Machine Learning, our first Kaggle submission
2	2022-02-02 Decision boundaries, <code>sklearn</code> pipelines, kNN classifier
3	2022-02-07 Validation, also quick intro to Pandas .
4	2022-02-09 Subsampling, Decision Trees
5	2022-02-14 Cross-validation, Bootstrap, Boosting, Bagging, Random Forests
6	2022-02-16 Grid search, multi-class classifiers
7	2022-02-23 Competition-specific analysis, PCA
8	2022-02-28 Stacking, Kernels, Support Vector Machines
9	2022-03-02 New competition, Intro to regression, parallelize, GLM and transformed targets
10	2022-03-07 Regularization, Feature Selection, Polynomial Regression
11	2022-03-09 Cross-validation for time-series
12	2022-03-14 Inner and Outer Database Joins
13	2022-03-16 Diagnostic plots with Yellowbrick
14	2022-03-21 Neural Networks
15	2022-03-23 Backpropagation, Convolutional Neural Networks
16	2022-03-28 Neural Networks with TensorFlow, Keras and TPU
17	2022-03-30 Dealing with overfitting
18	2022-04-04 Optimizers
19	2022-04-06 Transfer Learning
20	2022-04-11 Ethics
21	2022-04-13
22	2022-04-25 word2vec and autoencoders
23	2022-04-27 Recurrent neural networks
24	2022-05-02 Bayesian Statistics
25	2022-05-04 Bayesian Classifiers
26	2022-05-09 Reinforcement Learning
27	2022-05-11 Computational Creativity, Deep Dream, Style transfer, GAN
28	2022-05-16

Exam Review

Some exam review questions can be found [here](#)

Powered by [Pelican](#). Theme based on [Twitter Bootstrap](#).

© Mikael Vejdemo-Johansson