

Introduction to data science

Patrick Shafto

Department of Math and Computer Science

Plan for today

- Machine learning in SciKit Learn
- Homework (projects!)

Lecture

Machine Learning, Statistics and Probability

A tutorial introduction to machine learning with sklearn, an IPython-based slide deck by Andreas Mueller.

http://amueller.github.com/sklearn_tutorial

Spring course!

26:645:652 STATISTICS AND MACHINE LEARNING 3 credits Sections: 1 /							
SEC	INDEX	MEETING TIMES / LOCATIONS				EXAM	INSTRUCTORS
01 OPEN	14705	Monday	10:00 AM - 11:20 AM	NWK	ACK-123	A	YANG, C
		Wednesday	10:00 AM - 11:20 AM	NWK	ACK-123		

Enroll!

Monday!

Guest lecturer from industry

Please be on time

Come prepared with questions about data science in NYC

Projects!

First round due Nov 18th by 11:59 pm

Will then get comments

Give presentation in class

Turn in final assignment Dec 14 11:59 pm