

# SYS 4582/6018: Data Mining

Spring 2019

**Schedule**  
(./index.html)

**Syllabus**  
(./syllabus.html)

**Homework**  
(./homework.html)

## Schedule

Class	Date	Day	Topic	Reading (Pre-Class)	Notes
1	Jan 14	Mon	Introduction (lectures/00-intro.pdf)		Get textbook(s); install software
2	Jan 16	Wed	Association Analysis	MMDS 6.1; ITDM 5.1-5.2, 5.7	
	Jan 21	Mon	No Class: MLK Jr Day		
3	Jan 23	Wed	Association Analysis	ESL 14.2; MMDS 6.2; ITDM 5.3,5.8	
4	Jan 28	Mon	Networks		HW 1 Due
5	Jan 30	Wed	Networks		
6	Feb 04	Mon	Density Estimation		HW 2 Due
7	Feb 06	Wed	Density Estimation		
8	Feb 11	Mon	Clustering		HW 3 Due
9	Feb 13	Wed	Clustering		
10	Feb 18	Mon	Anomaly Detection		HW 4 Due
11	Feb	Wed	Anomaly Detection		

20				
12	Feb 25	Mon	Review	HW 5 Due
13	Feb 27	Wed	<b>Exam I</b>	
14	Mar 04	Mon	Supervised Learning	
15	Mar 06	Wed	Supervised Learning	
	Mar 11	Mon	No Class: Spring Break	
	Mar 13	Wed	No Class: Spring Break	
16	Mar 18	Mon	Resampling	HW 6 Due
17	Mar 20	Wed	Resampling	
18	Mar 25	Mon	Shrinkage and Model Selection	HW 7 Due
19	Mar 27	Wed	Shrinkage and Model Selection	
20	Apr 01	Mon	Classification	HW 8 Due
21	Apr 03	Wed	Classification	
22	Apr 08	Mon	Review	HW 9 Due
23	Apr 10	Wed	<b>Exam II</b>	
24	Apr 15	Mon	Tree-Based Methods	
25	Apr 17	Wed	Tree-Based Methods	

26	Apr 22	Mon	Boosting	HW 10 Due
27	Apr 24	Wed	Boosting	
28	Apr 29	Mon	Data Pre-Processing	
	May 06	Mon	Final Exam (2:00-5:00pm)	

## Reading

- **ISL: An Introduction to Statistical Learning** by James, Witten, Hastie and Tibshirani (<http://www-bcf.usc.edu/~garth/ISL/>)
- **ESL: The Elements of Statistical Learning: Data Mining, Inference, and Prediction (2nd Edition)** by Hastie, Tibshirani, and Friedman (<http://www-stat.stanford.edu/~tibs/ElemStatLearn/>)]
- **ITDM: Introduction to Data Mining (Second Edition)** by Tan, Steinbach, Karpadne, and Kumar (<https://www-users.cs.umn.edu/~kumar001/dmbook/index.php>)
- **MMDS: Mining of Massive Datasets** by Jure Leskovec, Anand Rajaraman, Jeff Ullman (<http://www.mmids.org/>)