

Schedule

IST 557 Data Mining: Techniques and Applications

Syllabus	Schedule	Assignments	Project
----------	----------	-------------	---------

- all homework/reports are due at midnight (11:59pm)
- schedule is subject to change
- black: lectures; blue: in-class discussions; orange: student presentation

Week	Date	Lecture (tentative schedule)	Slides	Deadline
1	8/27	Course overview Intro to data mining	0-class-intro.pdf	quiz quiz due 11:59pm 8/28 Wednesday
	8/29	Intro to classification	1-classification-intro.pdf	8/31 regular drop deadline 9/1 add deadline
2	9/3	Classification 1 - decision tree	2-decision-tree.pdf	
	9/5	Classification 1 - bagging & boosting	3-bagging.pdf 4-boosting.pdf	Assign 1
3	9/10	Assign 1 Q&A (TA in class)		Assign 1 due 11:59pm 9/10 Tuesday
	9/12	Classification 2 - perceptron, SVM	5-perception.pdf 6-svm.pdf	
4	9/17	Classification 2 - Kernel	7-kernel.pdf	Assign 2
	9/19	Classification 3 - bayesian statistics	8-bayesian.pdf	
5	9/24	Assign 2 Q&A (TA in class)		Assign 2 due 11:59pm 9/24 Tuesday
	9/26	Classification 4 - generative models logistic regression	9-generative-model.pdf 9-naive-bayes-example.p df 10-logistic.pdf	
6	10/1	Classification 5 - KNN	11-knn.pdf	
	10/3	Classification 6 - neural nets	12-neural-nets.pdf	Assign 3
7	10/8	Clustering 1	13-kmeans.pdf	Suggest your preferred Kaggle competition and preferred team members to TA (hzw77@psu.edu) by 10/8

	10/10	Assign 3 Q&A (TA in class)		Assign 3 due 11:59pm 10/10 Thursday
8	10/15	Clustering 2	14-mix.pdf 15-clustering.pdf	Project starts
	10/17	advanced topic - deep unsupervised learning	link to slides	
9	10/22	advanced topic - reinforcement learning	link to slides	
	10/24	Project Q&A (TA in class)		11:59pm 10/27 project check 1 (starting point) NO late submission. No extension will be granted to any team.
10	10/29	project presentation 1 (starting point)	A1 slides B1 slides C2 slides	
	10/31	Project Q&A (TA & top teams in class)		
11	11/5	invited lecture from industry (Facebook)		
	11/7	invited lecture from industry Hongjian Wang, Twitter		
12	11/12	advanced topic - representation learning	link to slides	
	11/14	Project Q&A (TA in class)		11/15 late drop deadline 11:59pm 11/17 project check 2 (midterm) NO late submission. No extension will be granted to any team.
13	11/19	project presentation 2 (midterm)	A6 slides B8 slides C1 slides	
	11/21	Project Q&A (TA & top teams in class)		
14	11/24-30	Thanksgiving week		
15	12/3	Data mining research		
	12/5	invited lecture from industry Alex Beutel, Google Research		11:59pm 12/8 project check 3 (final result) NO late submission. No extension will be granted to any team.
16	12/10	project check point 3 (final)		
	12/12	Report writing (no class, open office hour in E331 Westgate)		12/15 final report due NO late submission

