IST 5535 Schedule (Spring 2021)

Tentative Course Outline

The following is a general plan for the course; Deviations may be necessary.

1. **Introduction to Machine Learning** [Ch 1-2]

   Overview of machine learning; scales of measurement; assessing model accuracy.

2. **Getting Started with R**

   R platform, R Markdown, data structures, functions, control structures.

3. **Linear Regression** [Ch 3]

   Linear regression; other considerations of regression model.

4. **Classification** [Ch 4]

   Logistic regression; linear discriminant analysis; quadratic discriminant analysis.

5. **Resampling Methods** [Ch 5]

   Cross-validation; bootstrap.

6. **Linear Model Selection and Regularization** [Ch 6]

   Subset selection; shrinkage methods.

7. **Tree-Based Methods** [Ch 8]

   Regression and classification trees; ensemble learning and random forests; support vector machines; neural nets.

8. **Support Vector Machines** [Ch 9]

   Support vector classifiers; support vector machines.

9. **Unsupervised Machine Learning** [Ch 10]

   Principal component analysis; clustering; customer segmentation; association rules; market basket analysis.

10. **Neural Networks and Deep Learning** (if time allows)

    Neural nets; Keras; deep learning for text analytics.

Tentative Schedule

Zoom meeting time: TuTh 12:30 PM – 1:45 PM
Note: Deviations may be necessary according to class progress

Week	Topic	Due
W1: Jan 19, 21	Course Introduction; 1. Intro to Machine Learning	Form project group Due Jan 21
W2: Jan 26, 28	2. Getting Started with R *Topic Presentation #1 (Jan 28): LaValle, Lesser, Shockley, Hopkins and Kruschwitz [1]*	
W3: Feb 2, 4	3. Linear Regression *Topic Presentation #2 (Feb 4): Domingos [2]*	
W4: Feb 9, 11	4. Classification *Topic Presentation #3 (Feb 11): Ngai, Xiu and Chau [3]*	
W5: Feb 16, 18	5. Resampling Methods	
W6: Feb 23, 25	5. Resampling Methods	
W7: Mar 2, 4	6. Linear Model Selection and Regularization	
W8: Mar 9	No class on Mar 11 (Spring Recess)	Exam 1: Mar 9
W9: Mar 16, 18	7. Tree-Based Methods *Topic Presentation #4 (Feb 18):Ballings and Van den Poel [4]*	Project Deliverable 1 Due Mar 19
W10: Mar 23, 25	No class (Spring Break)	
W11: Mar 30, Apr 1	8. Support Vector Machines *Topic Presentation #5 (Apr 1):Veganzones and Séverin [5]*	
W12: Apr 6, 8	9. Unsupervised Machine Learning	
W13: Apr 13, 15	9. Unsupervised Machine Learning	Exam 2 on Apr 15
W14: Apr 20, 22	10. Neural Networks and Deep Learning	
W15: Apr 27; 29	10. Neural Networks and Deep Learning *Topic Presentation #6 (Jan 29):Du, Liu and Hu [6]*	
W16: May 4, 6	Wrap-up; Student Presentation	Project Deliverable 2: May 3
W17: May 11, 13	No Class	

Topic Presentations

Instructions:

1. Each presentation will be 15 minutes. Prepare an appropriate number of sides (neither too few nor too many).
2. Topic presentations by students need to cover at least the assigned paper(s). Find other relevant papers or resources to supplement your presentation whenever necessary.

Assigned Papers:

[1] LaValle, S., Lesser, E., Shockley, R., Hopkins, M.S., Kruschwitz, N., "Big data, analytics and the path from insights to value". MIT Sloan Management Review, 2011, pp. 21-32.

[2] Domingos, P., "A few useful things to know about machine learning". Communications of the ACM, 2012, pp. 78-87.

[3] Ngai, E.W.T., Xiu, L., Chau, D.C.K., "Application of data mining techniques in customer relationship management: A literature review and classification". Expert Systems with Applications, 2009, pp. 2592-2602.

[4] Ballings, M., Van den Poel, D., "Customer event history for churn prediction: How long is long enough?". Expert Systems with Applications, 2012, pp. 13517-13522.

[5] Veganzones, D., Séverin, E., "An investigation of bankruptcy prediction in imbalanced datasets". Decision Support Systems, 2018, pp. 111-124.

[6] Du, M., Liu, N., Hu, X., "Techniques for interpretable machine learning". Communications of the ACM, 2019, pp. 68-77.