INFO 329-03 Introduction to Data Mining

Lauren Laker
lakerlf@xavier.edu

Follow this and additional works at: https://www.exhibit.xavier.edu/management_information_systems_syllabi_fall_2017

Recommended Citation
https://www.exhibit.xavier.edu/management_information_systems_syllabi_fall_2017/18

This Restricted-Access Syllabus is brought to you for free and open access by the Management Information Systems Syllabi 2017 at Exhibit. It has been accepted for inclusion in Management Information Systems Syllabi Fall 2017 by an authorized administrator of Exhibit. For more information, please contact exhibit@xavier.edu.
INSTRUCTOR TO DATA MINING
INFO 329/MKTG 329
Fall 2017

Instructor: Dr. Lauren F. Laker
B.A. Mathematics & Statistics, Miami University
B.S. Secondary Math Education, Miami University
Masters of Business Administration, Xavier University
PhD Operations, Business Analytics, Information Systems, University of Cincinnati

Office Hours: T/TH 1:30-3:30 or by appointment
211 Smith Hall

Contact Info: 513.745.2034
lakerlf@xavier.edu

Authors: Ron Klimberg and B.D. McCullough
**MUST have SECOND edition**

Authors: Galit Shmueli, Peter Bruce, Mia Stephens, Nitin Patel

Optional for those in need of additional Stats help
Practical Data Analysis with JMP (ISBN: 978-1-61290-823-6)
Author: Robert Carver

Course Description:
Students will receive an introduction to the basic theory, tools and techniques of data mining, including prediction, associations, clustering, and recommendation systems. The course will be delivered from two points of view: the technological view and the marketing management view. Students will use data mining tools when doing their team projects.

Course Objectives:
In the last decade we have seen an explosion in the quantity of data available to businesses. Transactional data from point-of-sale scanners are now routinely available; data from direct marketing is growing exponentially; and e-commerce and web-browsing data is everywhere. Obviously, there is going to be a strong interest in extracting value or knowledge from this data. My vision of this course is to present and discuss data mining technologies and their application to data sets in an effort to support tactical and strategic business decisions. However, the over-riding focus will be learning when and how to use the technologies.

• Understand popular data mining techniques, how to apply them, and when they are applicable
• Utilize a state-of-the-art commercial data mining package
• Apply popular data mining techniques to analyze and solve real-world problems
• Recognize and assess ethical data mining practices

XAVIER UNIVERSITY VISION STATEMENT
"Xavier men and women become people of learning and reflection, integrity, and achievement, in solidarity for and with others."

WILLIAMS COLLEGE OF BUSINESS MISSION
"We educate students of business, enabling them to improve organizations and society, consistent with the Jesuit tradition."
Student Assessment and Evaluation
1) Class attendance, participation and discussion 5%
2) Exam 1 25%
3) Exam 2 25%
4) Course Project 10%
5) Class Readings/HW 30%
6) Final Exam 5%

Grade Distribution
93-100  A
90-92.9  A-
87-89.9  B+
83-86.9  B
80-82.9  B-
77-79.9  C+
73-76.9  C
70-72.9  C-
60-69.9  D
Below 60  F

COURSE POLICIES
1. Attendance and participation are critical in this course and will be recorded daily. This course is very interactive (hands-on)...Knowledge is CONSTRUCTED not received. I will take attendance every class period. This will be used as a measure for the In-Class portion of your grade. Attendance is not mandatory, but I WILL NOT COVER MISSED MATERIAL DURING OFFICE HOURS FOR THOSE WHO DO NOT ATTEND CLASS. If you must miss for a legitimate reason, please let me know and we can arrange a time to discuss what is/was missed.
2. No late work accepted. Assignments are to be submitted on the due date at the beginning of class. Late assignments will not be accepted unless PRIOR arrangements have been made with the instructor. A score of zero will be assigned for any assignment received beyond the due date.
3. No make-up exams/tests. They will be taken only those days in class. A score of zero will be assigned for any missed tests.
4. INFO/MKTG 329 has a zero tolerance for academic dishonesty.
5. Technology: The use of the computers during class should be used ONLY for the material. They are not to be used for any other purpose (email, instagram, etc.) during class time.

ACADEMIC HONESTY
“All work submitted for academic evaluation must be the student’s own. Certainly, the activities of other scholars will influence all students. However, the direct and unattributed use of another’s efforts is prohibited as is the use of any work untruthfully submitted as one’s own.” The penalty for violation of this policy will be a zero for that assignment if it is a first offense. Subsequent violation will result in an F for the course.

You are encouraged to work and study with other students in class and to learn from one another as opportunities provide. However, turning in the work of another or providing your work to someone else will be considered academically dishonest. It is my tendency to handle such cases with the severest penalties possible.
## Tentative Schedule

<table>
<thead>
<tr>
<th>Week of</th>
<th>Topic</th>
<th>Textbook</th>
</tr>
</thead>
<tbody>
<tr>
<td>8/22</td>
<td>Course Introduction</td>
<td>Klimberg: Chapter 1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Shmueli: Chapters 1, 2</td>
</tr>
<tr>
<td>8/29</td>
<td>Introduction to JMP &amp; Data Visualization</td>
<td>Klimberg: Chapter 4</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Shmueli: Chapters 2, 3</td>
</tr>
<tr>
<td>9/5</td>
<td>Finish Introduction to JMP &amp; Data Visualization</td>
<td>Klimberg: Chapters 2, 3</td>
</tr>
<tr>
<td></td>
<td>Start Statistics Review</td>
<td>Shmueli: Chapter 5</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Carver: Chapters 1, 3, 5, 9</td>
</tr>
<tr>
<td>9/12</td>
<td>Finish Statistics Review</td>
<td></td>
</tr>
<tr>
<td>9/19</td>
<td>Regression, ANOVA, Transformations</td>
<td>Klimberg: Chapter 5</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Shmueli: Chapter 6</td>
</tr>
<tr>
<td>9/26</td>
<td>Regression, ANOVA, Transformations</td>
<td>Carver: Chapters 14, 15</td>
</tr>
<tr>
<td>10/3</td>
<td>Regression, ANOVA, Transformations</td>
<td></td>
</tr>
<tr>
<td>10/10</td>
<td>EXAM #1 (2 part, both days)</td>
<td>Stats, Regression, ANOVA</td>
</tr>
<tr>
<td>10/17</td>
<td>Logistic Regression</td>
<td>Klimberg: Chapter 6</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Shmueli: Chapter 10</td>
</tr>
<tr>
<td>10/24</td>
<td>Decision Trees</td>
<td>Klimberg: Chapter 10</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Shmueli: Chapter 19</td>
</tr>
<tr>
<td>10/31</td>
<td>Neural Networks</td>
<td>Klimberg: Chapter 12</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Shmueli: Chapter 11</td>
</tr>
<tr>
<td>11/7</td>
<td>Model Comparison &amp; Assessment</td>
<td>Klimberg: Chapter 14</td>
</tr>
<tr>
<td>11/14</td>
<td>Exam #2 (2 part, both days)</td>
<td>Logistic Regression, Decision Trees, Neural Networks</td>
</tr>
<tr>
<td>11/21</td>
<td>Class on Tuesday but not Thursday THANKSGIVING - NO CLASS</td>
<td>Project Introduction</td>
</tr>
<tr>
<td>11/28</td>
<td>Cluster Analysis</td>
<td>Klimberg: Chapter 9</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Shmueli: Chapter 14</td>
</tr>
<tr>
<td>12/5</td>
<td>Final Exam &amp; Project Work</td>
<td></td>
</tr>
<tr>
<td>12/14</td>
<td>Group Project Presentations</td>
<td></td>
</tr>
</tbody>
</table>