IS7036: Data Mining for Business Intelligence

Spring 2018

February 26 – April 21, 2018
Section 1: Tuesday and Thursday: 4:00 – 5:50pm, Lindner 220
Section 2: Monday: 6:00 – 9:50pm, Lindner 107

Instructor: Dr. Zhe (Jay) Shan
Office: 317 Lindner Hall
Phone: 556-7006
Email: zhe.shan@uc.edu
Office Hours: Monday (3:00 – 5:00pm), and by appointment

Teaching Assistant: Mr. Tianhai Zu
Office: 534 Lindner Hall
Phone: 556-5703
Email: sundruri@mail.uc.edu
Office Hours: Friday (4:00 – 5:00pm), and by appointment

Course Overview & Objectives:
This course is designed for the in-depth learning of data mining knowledge and techniques in the context of business intelligence. The topics include association rules mining, classification, clustering, text mining & sentiment analysis, Web mining, and social analytics. Moreover, students will apply and integrate the business intelligence knowledge leant from IS7036 in IBM SPSS Modeler. This course also includes several case studies and a term project. The objectives of this class include:

- Understand the fundamentals of data mining
- Learn top data mining algorithms with their business applications
- Comprehend text mining and sentiment analysis
- Comprehend Web mining and analytics
- Practice those techniques in IBM SPSS Modeler
- Conduct a term project on data mining

Pre-requisite: IS7032 or IS6030 (If you don't have this pre-requisite, please contact instructor immediately.)

Required Textbook:
Ramesh Sharda, Dursun Delen, and Efraim Turban:
Business Intelligence and Analytics – Systems for Decision Support
Course Administration:
Class lectures will discuss extensively on the key concepts and ideas of each topic with additional insight. Most importantly, class lectures will clarify students' questions and initiate the lively discussion. The course involves lectures, 6 in-class lab sessions, 4 homework assignments, 1 midterm exam, and 1 final project.

I will create an active education environment to motivate you to understand and apply the information technologies for DM implementation. Therefore, attendance and class participation are critical. If your personal schedule entails missing some class sessions (more than 10 hours of classes), please drop this course. If you will be absent from a class, you have to notify the instructor by email in advance. Absence for a class does not relieve you of your responsibility for the subject matter, assignments when they are due, and other course-related issues discussed during that class period. Students that attend classes and participate in-class discussion often score much higher on exams, assignments and project.

Homework Assignments, Exam, and Final Project:
- There is one 120-minute midterm exam. The midterm contains both objective (e.g., T/F, multiple choices) and short-answer questions.
- There are no scheduled make-up tests, so it is important that you take all tests at their scheduled times. If you have a problem with a test date for a valid reason like a last minute accident or illness (job interviews don't qualify), you must contact the instructor immediately by email or phone before the actual test time, and present formal documents (e.g. doctor note, etc. Clinic visit summary does not qualify). Failure to do so will mean getting a zero score.
- Late submissions of homework cannot be accepted unless there is a real emergency (job interviews don't qualify) AND prior permission of instructor has been obtained.
- You may appeal any score on a test or an assignment, but such appeal must be made in writing or in person within one week from the time the test or homework is returned to the class.
- The details of the final project will be announced in week 3.

Course Grade:
Your final course grade will be determined as:
1. Homework Assignments  20%
2. Midterm Exam        40%
3. Final Project        40%
Total                      100%
Evaluation:
The final grade will be based on the percentage of the total possible points awarded as follows:

<table>
<thead>
<tr>
<th>GRADE</th>
<th>Scores</th>
<th>GRADE</th>
<th>Scores</th>
<th>GRADE</th>
<th>Scores</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>100-93</td>
<td>B</td>
<td>82-79</td>
<td>C</td>
<td>68-64</td>
</tr>
<tr>
<td>A-</td>
<td>92-88</td>
<td>B-</td>
<td>78-74</td>
<td>C-</td>
<td>63-60</td>
</tr>
<tr>
<td>B+</td>
<td>87-83</td>
<td>C+</td>
<td>73-69</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Communication:
Email is the best way to communicate with me. I intend to reply your emails within 24 hours (48 hours in weekend). Therefore, I expect the same courtesy from you. Email will be a very important mode of communication in this class as we meet only once/twice a week. Please start your subject line of your emails with ‘IS7036-00X’ (00X is your section number: 001 for the afternoon section, and 002 for the evening section).

Also make sure to check Blackboard on a regular basis for updates or additional course material. Please keep your email information on Blackboard up to date. Whenever possible, I will announce changes to the course via the Blackboard announcement function.

Class Conduct:

**IMPORTANT !**
The use of any mobile device, such as laptop, smart phones, etc., during a lecture session is not permitted unless the instructor has approved its use. Please silence your devices and put them in your bag or jacket. Please be professional!

One of our goals in the College of Business is to prepare you for a professional career and we wish to instill good habits through insisting on a high level of conduct in all your classes. The following behaviors are unacceptable in this class:

1. Not on time
2. Failing to notify the professor by email when you are absent from class.
3. Excessive tardiness/absences.
4. Leaving while class is in session for reasons other than an emergency.
5. Working on non-course related assignments.
6. Falling asleep or resting head on desk.
7. Wearing headphones while class is in session.
8. Web surfing in class
9. Other types of disrespectful behaviors.

Academic Integrity:

**IMPORTANT !**
I personally take academic integrity very seriously, and if there are any violations, the consequences will be unfortunate. Please don't take this issue lightly because plagiarism in any form will not be tolerated.

As with all Lindner College of Business efforts, this course will uphold the highest ethical standards, critical to building character (the C in PACE). Ensuring your integrity is vital and your responsibility. LCB instructors are required to report ANY incident of academic misconduct (e.g., cheating, plagiarism) to the college review process, which could result in severe consequences, including potential dismissal from the college. For further information on Academic Misconduct or related university policies and procedures, please see the UC Code of Conduct (http://www.uc.edu/conduct/Code_of_Conduct.html).

Special Needs:
If you have any special needs related to your participation in this class that may influence your performance you should inform and meet the instructor immediately to arrange reasonable provisions to ensure an equitable opportunity to meet all the class requirements.

**Weather-Related Policy:**

When inclement weather threatens the safety of the University of Cincinnati community, the Senior Vice President for Administration and Finance may invoke University Rule 3361: 10-55-01 and declare an emergency closing. It is possible that only some campuses of the university may close. Please sign our related alert messages on [http://www.uc.edu/alert.html](http://www.uc.edu/alert.html) for latest updates.

If campus closure occurs, lectures will be delivered in online platform at same time slots via Web conference tools (e.g. WebEx). But missed exams will be rescheduled and related details will be announced in email. In cases where an exam make-up session is necessary and students have work commitments or travel arrangements that cannot easily be changed, the student should work with instructors to find a means of accommodation. All assignments are submitted electronically on Blackboard, and therefore their due dates will not be altered, unless other special notices.

Students are expected to make reasonable accommodations in work schedules or travel plans to accommodate weather-related changes. However, instructors will give due consideration to “impossibilities to perform” which might include (but is not limited to): 1. Students without internet access at home; 2. Failures of technology (internet or power); 3. Need for special software or equipment available on campus; 4. Local emergencies that do not close the UC campus but might prevent a student from coming to campus. The student should work with instructors to find a reasonable way to make up missed material.

Closure of the University does not relieve students of the responsibility for completing all required course work. Within practical limits, work should be completed in a timely manner and within stated deadlines. Students are expected to have a valid email address accessible to the instructor through courseware (Blackboard). Students are expected to check email and/or other means of electronic communication for information from the instructor during a closing.

If the University is open, but a student is unable to attend a class or activity due to an emergency declared for their area the student should inform the instructor in a timely manner and pursue appropriate make-up opportunities.
Course Schedule (Tentative):

<table>
<thead>
<tr>
<th>Module</th>
<th>Topics</th>
<th>Related Chapters</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>Course Introduction and Logistics</td>
<td>Syllabus</td>
</tr>
<tr>
<td>1</td>
<td>Intro to Data Mining</td>
<td>Sharda: Chapter 5</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Handouts</td>
</tr>
<tr>
<td>2</td>
<td>Association Rules</td>
<td>Sharda: Chapter 5</td>
</tr>
<tr>
<td></td>
<td><em>SPSS Lab 1. Aprior</em></td>
<td>Handouts</td>
</tr>
<tr>
<td>3</td>
<td>Classification</td>
<td>Sharda: Chapter 5</td>
</tr>
<tr>
<td></td>
<td><em>SPSS Lab 2. CHAID</em></td>
<td>Handouts</td>
</tr>
<tr>
<td>4</td>
<td>Advanced Classification</td>
<td>Sharda: Chapter 6</td>
</tr>
<tr>
<td></td>
<td><em>SPSS Lab 3. ANN</em></td>
<td>Handouts</td>
</tr>
<tr>
<td>5</td>
<td>Clustering</td>
<td>Sharda: Chapter 5</td>
</tr>
<tr>
<td></td>
<td><em>SPSS Lab 4. K-means</em></td>
<td>Handouts</td>
</tr>
<tr>
<td>6</td>
<td>Web Mining</td>
<td>Sharda: Chapter 8</td>
</tr>
<tr>
<td></td>
<td><em>SPSS Lab 5. Web Mining</em></td>
<td>Handouts</td>
</tr>
<tr>
<td>7</td>
<td>Text Mining</td>
<td>Sharda: Chapter 7</td>
</tr>
<tr>
<td></td>
<td><em>SPSS Lab 6. Text Mining</em></td>
<td>Handouts</td>
</tr>
</tbody>
</table>