MBA 738: Data Mining for Business Intelligence (3 credits)

SPRING 2015

1. Meeting Time and Place
    Location: Founders Hall, Room 118
    Time: Wednesdays 6:30 pm - 10:05 pm

2. Instructor Information
    Name: Pallab Sanyal, Ph. D.
    Email: psanyal@gmu.edu
    Office: Enterprise Hall, Room 148
    Office Phone: (703) 993-1888
    Office Hours: Wednesdays 5:00-6:00 PM, or by appointment.

3. Course Materials
   a. Optional Textbooks:
         The book can be downloaded for free from the Global Text Project website (link), but if you prefer a bound copy, you can purchase it from online retailers such as Amazon.com (link).

   b. Required Software: RapidMiner Studio.
      The Starter edition of the software can be downloaded for free for Windows, Mac and Linux systems. Registration and installation instructions are provided on Blackboard.

   c. Articles and Cases: All articles and cases for this course can be downloaded for free from the Web or the GMU library (http://library.gmu.edu). The links will be provided on Blackboard.

4. Course Description
   Data mining—the art of extracting useful information from large amounts of data—is of growing importance in today’s world. The amount of data flowing from, to, and through enterprises of all sorts is enormous, and growing rapidly—more rapidly than the capabilities of organizations to use it. Enterprises are trying to make effective use of the abundance of data to which they have access:
to make better predictions, better decisions, and better strategies. Therefore, managers now need to know about the possibilities and limitations of data mining. This course will provide an introduction to data mining problems and tools to enhance managerial decision making. The students will learn how to ask the right questions and how to draw inferences from the data by using the appropriate data mining tools. Using the RapidMiner software, the students will acquire hands-on experience on applying data mining methods. Overall, the course will enable students to approach business problems data-analytically, envision data-mining opportunities in organizations, and also follow up on ideas or opportunities that present themselves.

5. Course Objectives
   a. Introduce the steps involved in data mining, from goal definition to model deployment.
   b. Discuss data preparation techniques.
   c. Examine supervised learning methods, such as classification and prediction.
   d. Examine unsupervised learning methods, such as association rules and clustering.

6. Grading and Assessment

   Grade Distribution

<table>
<thead>
<tr>
<th>Grade</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>A/A-</td>
<td>&gt;= 90%</td>
</tr>
<tr>
<td>B+/B/B-</td>
<td>80% to &lt; 90%</td>
</tr>
<tr>
<td>C</td>
<td>70% to &lt; 80%</td>
</tr>
<tr>
<td>D</td>
<td>60% to &lt; 70%</td>
</tr>
<tr>
<td>F</td>
<td>below 60%</td>
</tr>
</tbody>
</table>

   Split between +/- scores will be determined by the instructor based on clustering of scores.

   Instructions for all assignments will be posted on Blackboard. Completed written assignments should be submitted via Blackboard only.

   **Class Participation:** You will be expected to participate in class discussions and complete in-class exercises. Each week we will discuss relevant articles and/or work through small exercises in class. You will be evaluated based on your involvement in these and other discussions in class. You are encouraged to discuss your own work experience when relevant to the material being covered in class. You are also encouraged to ask questions in class.

   The following factors will contribute positively to your participation score: (i) Arriving before the start of class and staying till the end, (ii) Listening actively to the instructor and peers, (iii) Asking good questions, (iv) Responding to questions asked to the class, (v) Neither dominating the conversation nor being too quiet, and (vi) Exhibiting a good sense of humor.

   The following factors will contribute negatively to your participation score: (i) Arriving after the start of the class and/or leaving before the end, (ii) Lack of involvement, silence, detachment or disinterest, (iii) Distracting others by surfing the web, e-mailing, texting (iv) Not listening actively, and (v) Leading the discussion into unrelated topics.

   10% of the class participation points will be awarded for submitting your resume before the first class.

   **Project:** The purpose of the project will be for you to gain hands-on experience in solving a realistic problem using the data mining principles covered in class. This will be a *group assignment*. You will be required to submit a report by the time and date designated in the schedule.
Homework: These will consist of primarily problem sets that are designed to give you valuable practice and enhance your understanding of the concepts covered in class. These will also be group assignments due by the times and dates designated in the schedule. If you submit all the assignments, your lowest score will be dropped from the final grade calculation.

Final Exam: This will be an individual take-home assignment due by the time and date designated in the schedule.

Semester Grade: Your semester grade will be assigned based on the total points earned on the assignments described above; no extra credit will be available. A solid job on all the assignments will be evaluated at the A-/B+ border. To earn an A, performance must go beyond “meets expectations.” You can request a review of any grade within a week following the assignment of grades. After that period no grade will be revised.

7. Academic Integrity
GMU is an Honor Code university; all students are responsible for knowing and following the GMU Honor Code Statement: “Student members of the George Mason University community pledge not to cheat, plagiarize, steal, or lie in matters related to academic work.” In the event of a violation of the GMU Honor Code, the violating student will be reported to the GMU Honor Committee. Another aspect of academic integrity is the free play of ideas. Discussions are encouraged in this course, with the firm expectation that all aspects of the class will be conducted with civility and respect for differing ideas, perspectives, and traditions. Please refer to http://oai.gmu.edu for further details. When in doubt (of any kind), please ask the instructor for guidance and clarification.

8. Learning Goals
MBA Program Learning Goals:
   a. Teaming & Leading: Our graduates will demonstrate the team leadership and interpersonal skills needed to form, lead, and work effectively on diverse organizational teams.
   b. Analytical Decision Making: Our students will demonstrate the ability to analyze uncertain complex management situations using appropriate tools, techniques and information systems for decision-making.
   c. Knowledge of Functional Business Disciplines: Our graduates will demonstrate the ability to integrate knowledge from all functional areas of business into a meaningful firm-level perspective.
   d. Global Understanding: Our graduates will demonstrate a perspective on how businesses operate in the global environment.
   e. Communication Skills: Our graduates will demonstrate written, oral and presentations skills necessary to explain problems and solutions effectively and persuasively.
   f. Ethics and Social Responsibility: Our graduates will have a sense of professional and social responsibility in the conduct of managerial affairs.

9. Learning Disabilities
If you are a student with a disability and you need academic accommodations, please see me and contact the Disability Resource Center (DRC) at 703-993-2474, at the beginning of the semester. All academic accommodations must be arranged through the DRC.
10. Other Course Policies

a. **Attendance:** Attendance in class is mandatory. If you are absent, it is your responsibility to find out from a classmate what you missed (both course materials and announcements).

b. **E-Mail Correspondence:** Outside of the designated class time and office hours, e-mail is the easiest and quickest method to contact me. Consistent with federal privacy laws, I do not respond to non-GMU email ids with confidential information. *Any email that you send me must include "MBA 738" in the subject and your full name in the body.*

c. **Laptops and hand-held devices:** Technology can greatly assist learning, but it can also be a distraction. Laptops or any other hand-held devices should strictly be used for class related activities such as working on in-class exercises, taking notes or following lecture slides.
### Course Schedule *

<table>
<thead>
<tr>
<th>Week</th>
<th>Date</th>
<th>Topics</th>
<th>Chapters in Textbooks</th>
<th>Assignments due**</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>02/04</td>
<td>Overview of Data Mining</td>
<td>North: Ch. 1, 2 &amp; 3 Linoff &amp; Berry: Ch. 1 &amp; 3</td>
<td>HW 0 (Resume)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Overview of Data Mining; Data Preprocessing</td>
<td>North: Ch. 1, 2 &amp; 3 Linoff &amp; Berry: Ch. 1 &amp; 3</td>
<td>Groups for Homework</td>
</tr>
<tr>
<td>2.</td>
<td>02/11</td>
<td>Prediction Models</td>
<td>North: Ch. 8 Linoff &amp; Berry: Ch. 5 &amp; 6</td>
<td>HW 1</td>
</tr>
<tr>
<td>3.</td>
<td>02/18</td>
<td>Prediction Models</td>
<td>North: Ch. 8 Linoff &amp; Berry: Ch. 5 &amp; 6</td>
<td></td>
</tr>
<tr>
<td>4.</td>
<td>02/25</td>
<td>Prediction Models</td>
<td>North: Ch. 8 Linoff &amp; Berry: Ch. 5 &amp; 6</td>
<td>Groups for Project</td>
</tr>
<tr>
<td>5.</td>
<td>03/04</td>
<td>Classification Models</td>
<td>North: Ch. 9 Linoff &amp; Berry: Ch. 5 &amp; 6</td>
<td>HW 2</td>
</tr>
<tr>
<td>6.</td>
<td>03/11</td>
<td>Classification Models</td>
<td>North: Ch. 9 Linoff &amp; Berry: Ch. 5 &amp; 6</td>
<td></td>
</tr>
<tr>
<td>7.</td>
<td>03/18</td>
<td>Decision Tree Models</td>
<td>North: Ch. 10 Linoff &amp; Berry: Ch. 7</td>
<td>HW 3; Project Preview</td>
</tr>
<tr>
<td>8.</td>
<td>03/25</td>
<td>Cluster Analysis</td>
<td>North: Ch. 6 Linoff &amp; Berry: Ch. 12 &amp; 13</td>
<td>HW 4</td>
</tr>
<tr>
<td>9.</td>
<td>04/01</td>
<td>Cluster Analysis</td>
<td>North: Ch. 6 Linoff &amp; Berry: Ch. 12 &amp; 13</td>
<td></td>
</tr>
<tr>
<td>10.</td>
<td>04/08</td>
<td>Affinity Analysis; Review</td>
<td>North: Ch. 5 Linoff &amp; Berry: Ch. 15</td>
<td>HW 5</td>
</tr>
<tr>
<td></td>
<td>04/10</td>
<td></td>
<td></td>
<td>Project Report</td>
</tr>
<tr>
<td></td>
<td>04/11</td>
<td></td>
<td></td>
<td>Final Exam</td>
</tr>
</tbody>
</table>

*The schedule is tentative and subject to change.

** All assignments are due by 6 pm on the designated dates.