Course title and number: Data Mining and Business Analytics INFS 8348-01
Term: Fall 2018
Meeting Times: M 09:20 AM -12:00 PM Wes Cntr for Innov & Commercia 2.206
Instructor Name: Dr. Francis Kofi Andoh-Baidoo
Email: Francis.AndohBaidoo@utrgv.edu
Telephone: (956) 665-3397 Fax: (956) 665-3367
Office location: MAGC 3.322
Office Hours: Mondays: 1PM-3PM, Tuesdays & Thursdays: 11:00 AM – 12:00 PM

Course Description
In recent years, modern enterprises in the public and private sectors and academia have come to realize the importance of using integrated and diverse sources of data both within and without to deliver necessary information to appropriate individual or group to ensure effective and efficient decision making in this competitive era. Decline in computer processing as well as storage media costs have made it easy for organizations to store large amount of data from existing transactional processing systems. Data Mining and Business Analytics are novel tools and concepts that organizations employ to deliver business intelligence. In this course we will explore advanced concepts related to data mining & business analytics and their applications in modern organizations. Material for the course will be obtained from various sources including published research papers. A high level of student participation is required for successful completion of this course, particularly in preparation of presentation materials, preparation of review reports and in writing of synthesis papers.

Required Textbook:

No Required Textbook. Relevant material will be from selected published research papers.

Use of Web Site

Course materials are posted online on Blackboard via https://my.utrgv.edu/. For questions and technical support on using Blackboard, please contact the Center for Online Learning, Teaching and Technology (COLT) help desk at (956) 665-2927.

Course Learning Goals and Outcomes

The major objectives of this course are:
1. To provide students with the opportunity for gaining a deep understanding of selected advanced concepts of data mining & business analytics and their applications in modern organizations.
2. To provide students with the opportunity to gain knowledge on contemporary research issues in data mining & business analytics.
3. To familiarize students with techniques used to conduct research in data mining & business analytics.
4. To develop each student’s research skills in data mining & business analytics.
Robert C. Vackar College of Business and Entrepreneurship Learning Goals and Mission

The Robert C. Vackar College of Business and Entrepreneurship has adopted the following learning goals and objectives that each student should achieve while in the degree program. Not all goals and objectives are covered in each course; however, all goals and objectives should be addressed throughout the degree program.

<table>
<thead>
<tr>
<th>PhD Learning Goals</th>
<th>This course contributes to the following College of Business and Entrepreneurship learning goals:</th>
<th>How measured</th>
</tr>
</thead>
<tbody>
<tr>
<td>Goal 1: The acquisition of advanced knowledge in area of specialization.</td>
<td>X</td>
<td>Assignments, Project</td>
</tr>
<tr>
<td>Goal 2: The development of advanced theoretical or practical research skills for the area of specialization.</td>
<td>X</td>
<td>Assignments, Project</td>
</tr>
<tr>
<td>Goal 3: Preparation for teaching responsibilities in higher education, especially for those students who expect to enter teaching careers.</td>
<td>X</td>
<td>Assignments, Project</td>
</tr>
<tr>
<td>Goal 4: Dissertation, or equivalent, demonstrating personal integration of, and original intellectual contribution to the field of knowledge.</td>
<td>X</td>
<td>Assignments, Project</td>
</tr>
</tbody>
</table>

Calendar of Activities

Important dates for fall 2018 include:

- August 27: First day of classes
- August 30: Last day to add a course or register for fall 2018
- September 3: Labor Day – NO classes
- November 14: Last day to drop a course; will count toward the 6-drop rule
- November 22 - 24: Thanksgiving Holiday – NO classes
- December 6: Study Day – NO classes
- December 7 - 13: Final Exams
- December 14 – 15: Commencement Exercises

STUDENTS WITH DISABILITIES

If you have a documented disability (physical, psychological, learning, or other disability which affects your academic performance) and would like to receive academic accommodations, please inform your instructor and contact Student Accessibility Services to schedule an appointment to initiate services. It is recommended that you schedule an appointment with Student Accessibility Services before classes start. However, accommodations can
be provided at any time. **Brownsville Campus:** Student Accessibility Services is located in Cortez Hall Room 129 and can be contacted by phone at (956) 882-7374 (Voice) or via email at ability@utrgv.edu. **Edinburg Campus:** Student Accessibility Services is located in 108 University Center and can be contacted by phone at (956) 665-7005 (Voice), (956) 665-3840 (Fax), or via email at ability@utrgv.edu.

**MANDATORY COURSE EVALUATION PERIOD**

Students are required to complete an ONLINE evaluation of this course, accessed through your UTRGV account ([http://my.utrgv.edu](http://my.utrgv.edu)); you will be contacted through email with further instructions. Students who complete their evaluations will have priority access to their grades. Online evaluations will be available on or about **November 15 – December 5** for full fall semester courses.

**ATTENDANCE**

Students are expected to attend all scheduled classes and may be dropped from the course for excessive absences. UTRGV’s attendance policy excuses students from attending class if they are participating in officially sponsored university activities, such as athletics; for observance of religious holy days; or for military service. Students should contact the instructor in advance of the excused absence and arrange to make up missed work or examinations.

**SCHOLASTIC INTEGRITY**

As members of a community dedicated to Honesty, Integrity and Respect, students are reminded that those who engage in scholastic dishonesty are subject to disciplinary penalties, including the possibility of failure in the course and expulsion from the University. Scholastic dishonesty includes but is not limited to: cheating, plagiarism, and collusion; submission for credit of any work or materials that are attributable in whole or in part to another person; taking an examination for another person; any act designed to give unfair advantage to a student; or the attempt to commit such acts. Since scholastic dishonesty harms the individual, all students and the integrity of the University, policies on scholastic dishonesty will be strictly enforced (Board of Regents Rules and Regulations and UTRGV Academic Integrity Guidelines). All scholastic dishonesty incidents will be reported to the Dean of Students.

**SEXUAL HARASSMENT, DISCRIMINATION, and VIOLENCE**

In accordance with UT System regulations, your instructor is a “responsible employee” for reporting purposes under Title IX regulations and so must report any instance, occurring during a student’s time in college, of sexual assault, stalking, dating violence, domestic violence, or sexual harassment about which she/he becomes aware during this course through writing, discussion, or personal disclosure. More information can be found at [www.utrgv.edu/equity](http://www.utrgv.edu/equity), including confidential resources available on campus. The faculty and staff of UTRGV actively strive to provide a learning, working, and living environment that promotes personal integrity, civility, and mutual respect in an environment free from sexual misconduct and discrimination.
COURSE DROPS

According to UTRGV policy, students may drop any class without penalty earning a grade of DR until the official drop date. Following that date, students must be assigned a letter grade and can no longer drop the class. Students considering dropping the class should be aware of the “3-peat rule” and the “6-drop” rule so they can recognize how dropped classes may affect their academic success. The 6-drop rule refers to Texas law that dictates that undergraduate students may not drop more than six courses during their undergraduate career. Courses dropped at other Texas public higher education institutions will count toward the six-course drop limit. The 3-peat rule refers to additional fees charged to students who take the same class for the third time.

Email Policy

I check my Blackboard Mail regularly and will reply your email within 24 hours during the weekdays. This is the preferred communication medium between students and the instructor.

ATTENDANCE

Students are expected to attend all scheduled classes and may be dropped from the course for excessive absences. UTRGV’s attendance policy excuses students from attending class if they are participating in officially sponsored university activities, such as athletics; for observance of religious holy days; or for military service. Students should contact the instructor in advance of the excused absence and arrange to make up missed work or examinations.

Topic Areas:
1. Data Mining
2. Business Analytics
3. Classification
4. Clustering
5. Association Rule
6. Text Analytics
7. Big Data Analytics
8. Business Intelligence

Components of Pedagogical Approach
- Class Presentations & Discussions
- Preparation of Review Reports of Selected Research Papers
- Preparation of Synthesis Papers
- Preparation of Development of a Critical Bibliography

Class Presentations and Discussions
- Each paper will lead two seminars: (i) individual, and (ii) in a team of two or more students. (See Appendix 1).
- Presentations will be structured based on the guidelines in Appendix 2.
- Each student must contribute to the class discussion on each research paper. The instructor will randomly select students to contribute to the class discussion. If at anytime a selected student does not do a credible job of discussing the topic of the day, that student will lose 50% of his/her participation grade.
**Review Reports of Research Paper:**
For each research paper, a report containing a critical review and analysis of the relevant material will be prepared by each student or a team of students. Each student will prepare two (2) Review Reports. *See Appendix 2* for relevant guidelines. In addition, each student will prepare two (2) to three (3) page summary on each paper that is not assigned to the student or his/her team.

**Synthesis Papers:**
Synthesis papers involve synthesizing material from multiple design science research papers (*See Appendix 3*). In some cases, the goal will be to define a Data Mining or Business Analytics research problem and to explore the design of solution alternatives.
- Each student will prepare two (2) Synthesis Papers.
- The objective(s) and primary research papers that are associated with each Synthesis Paper would be discussed between students and the instructor.

**Critical Bibliography:**
Each student or team will prepare a critical bibliography on data mining & business analytics research. This critical bibliography will less depth than the review reports and synthesis paper. It is recommended that for each class session you should add papers to your critical bibliography that are relevant to the Research Papers of that session. *See Appendix 4* for relevant guidelines.

**Term Project:**
Each student will complete a term project based on the synthesis papers or a design science approach to addressing a research problem. The completed project should be targeted at ICIS 2019. The completed project should follow the design-oriented approach that will be discussed throughout the semester.

**Grading Scheme**
<table>
<thead>
<tr>
<th>Component</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Class Participation</td>
<td>10%</td>
</tr>
<tr>
<td>Review Report</td>
<td>20%</td>
</tr>
<tr>
<td>Synthesis Paper</td>
<td>20%</td>
</tr>
<tr>
<td>Critical Bibliography</td>
<td>20%</td>
</tr>
<tr>
<td>Term Project</td>
<td>30%</td>
</tr>
</tbody>
</table>
## APPENDIX 1:
### Tentative Schedule

<table>
<thead>
<tr>
<th>Session</th>
<th>Papers</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Aug 27</td>
<td>Introduction and Course Overview</td>
</tr>
<tr>
<td>2 Sep 03</td>
<td>Labor Day, No Classes</td>
</tr>
<tr>
<td>8 Oct 15</td>
<td><em><strong>Presentation of Term Project Proposal</strong></em></td>
</tr>
<tr>
<td>Date</td>
<td>Event Description</td>
</tr>
<tr>
<td>--------</td>
<td>-------------------</td>
</tr>
<tr>
<td>Oct 22</td>
<td><em><strong>Students work on Term Project</strong></em></td>
</tr>
<tr>
<td>Oct 29</td>
<td><em><strong>Students work on Term Project</strong></em></td>
</tr>
</tbody>
</table>
| Nov 05 | Association Rules  
| Nov 12 | **Big Data Analytics**  
| Nov 19 | **Business Intelligence**  
*** Students meet with Instructor to discuss Term Project *** |
| Nov 26 | *** Students meet with Instructor to discuss Term Project *** |
| Dec 03 | ***Students Presentation on Term Project*** |
| Dec 10 | ***No Final Exam*** |

B: Behavioural/Managerial/Empirical  
D: Design Science
APPENDIX 2:
Guidelines for Preparing Review Reports of Selected Research Papers

The length of your Review Report should be between approximately 17-10 pages. At a minimum a student or a team needs to do the following:

**Design Science (DS) Papers:**
1) Provide an Informal Description of the problem that the paper seeks to address. *This description of the problem is relatively general. You need to demonstrate that you understand this research problem.*

2) Provide a Formal Definition of the Research Problem: In many cases the research problem that is actually addressed by the authors is a restricted version of the informal problem description that they provided in the abstract and/or introduction section. Your discussion on the formal definition of the research problem should include:
   a. A narrative that demonstrates that you have an in-depth understanding of the formal research problem that is actually addressed in the paper. This should include:
      i. Identification & description of all relevant issues that are covered (i.e. issues explicitly or implicitly raised by authors; & weaknesses in the works of other researchers that were raised by authors)
      ii. Identification & description of all relevant issues that are not covered (i.e. issues explicitly excluded by authors; other issues relevant to the Informal Research problem that are not explicitly or implicitly covered)
      iii. Description of all assumptions and limitations.
   b. Definitions of all concepts & terms that the author(s) used in their conceptual framing of the research problem
   c. Discuss how the authors justify their conceptual framing of the research problem, including assumptions
   d. Include comments on both the weaknesses and strengths of the author’s conceptual framing of the research problem.

   Please note that in some cases material for the formal description of the research problem may be at various locations in the paper, depending on the author’s presentation approach.

3) Discuss how the authors establish the importance of their version of the given research problem.

4) Provide an in-depth description of the solution, as well as the authors’ justification of the solution. This would include describing how each issue covered in the formal definition of the research problem (see sub-sub-step 2aii above). You need to demonstrate that you understand the solution for addressing this research problem that is presented in the paper.

5) Discuss strengths and limitations of the solution that is presented. *In assessing the strengths and limitations of the solution, you should focus on the formal description of the research problem, as this is the problem that the authors promised to address. The question here is: Did the solution effectively and/or efficiently address this restricted problem (i.e. issue covered in the formal definition of the research problem)?*

6) Discuss the evaluation approach taken for establishing the effectiveness/efficiency of the solution. Discuss the results of the evaluation, and the strengths and limitations of the evaluation approach.

7) Suggest how the limitations of the solution could be addressed and how the content of the research could be otherwise improved

8) Identify the skills, knowledge and resources that are needed to adequately address this research problem.

**Behavioural/Managerial/Empirical (B) Papers:**

1) Provide an Informal Description of the problem that the paper seeks to address. Similar to (1) for Design Science (DS) papers.

2) Provide a Formal Definition of the Research Problem. Similar to (2) for DS papers.

3) Discuss how the authors establish the importance of their version of the given research problem. Similar to (3) for DS papers.
4) Provide an in-depth description of the research methodology, the authors’ justification of their research methodology.
5) Present & discuss the results
6) Suggest how the content of the research could be improved.
7) Similar to (8) for DS papers.
APPENDIX 3: Structure for Synthesis Paper(s)

1) Table of Contents

2) Description of Resources Papers:

<table>
<thead>
<tr>
<th>Paper Title</th>
<th>Role</th>
<th>Formal Definition of Research Problem</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Anchor</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Support</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Support</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Support</td>
<td></td>
</tr>
</tbody>
</table>

**Anchor**: One of the design science research papers covered in the class sessions. Your *New Research Problem* is an extension of the formal research problem that was covered in this paper.

**Support**: At a minimum, this includes three other design science research papers that were covered in the class sessions.

3) Provide a **Formal Definition** of your *New Research Problem*. Your new research problem should be an extension of the formal research problem of the anchor paper that also includes problem specific issues from all the support papers. This Formal Definition of the research problem has some similarity to the Formal Requirements Specification that is associated with the systems development process. Your discussion on the formal definition of the research problem should include:
   a. A narrative that would facilitate the reader’s attempt to obtain an in-depth understanding of the formal research problem that is actually addressed in the paper. This should include:
      i. Identification & description of all relevant issues that are covered
      ii. Identification & description of all relevant issues that are not covered
      iii. Description of all assumptions and limitations.
   b. Definitions of all concepts & terms used in your formal definition of your *New Research Problem*.
   c. Present your justification for your formal definition of your *New Research Problem*, including assumptions.
   d. Provide comments on both the weaknesses and strengths of the formal definition of your *New Research Problem*.

4) Provide a discussion that establishes the importance of your *New Research Problem*.

5) a. Provide a description of your **Solution Approach** for your *New Research Problem*, as well as your justification of the solution approach. At a minimum, you need to demonstrate that your solution approach adequately addresses issues identified in 3a-i above.
   b. Provide information on method used to collect data and the type of data mining or data analytics technique(s) employed.

6) Discuss strengths and limitations of your **Solution Approach**.

7) Discuss the **Evaluation Approach** that you would take to establish the effectiveness/efficiency of the solution approach. Provide justification for your evaluation approach. Discuss the strengths and limitations of the propose evaluation approach.

8) Suggest how the limitations of the solution approach could be addressed.

9) Identify the skills, knowledge and resources that are needed to adequately address this research problem

10) **List of References**
APPENDIX 4: Guidelines for Preparing Critical Biography

A. Topic Areas:
- See main body of course outline
- For each topic area, include both seminal and contemporary papers

B. Material to be included for each paper:
- Topic Area
- Paper Title
- Referencing data (i.e. author(s), journal/conference/workshop)
- Abstract
- Orientation: Technological, Managerial, Behavioral, Theoretical, Empirical, etc., or any combination of these.
- Comments on the Strengths and Weaknesses/Limitations of the paper. Do not focus on the presentation quality but rather on the content.
- Identify the Skills, Knowledge, and Resources that are needed to adequately address this research problem.