Semester: Spring 2014

Course Prefix/Number: CAP5771

Course Title: Data Mining

Course Credit Hours: 3.0

Course Meeting Times/Places:
Online

Instructor and Contact Information:
Dr. Dallas Snider
Office: Building 4, Room 233
E-mail: dsnider@uwf.edu
Tel: 850-473-7348

Office Hours: Monday, Wednesday and Friday 10:00am – noon
Other days and times by appointment

Course Web Site: http://elearning.uwf.edu/ (login and select Data Mining, CAP5771)

Prerequisites or Co-requisites: Database Systems (COP4710 or COP5725).

Course Description:
This course will expose students to data mining concepts, techniques, and software utilized in the overall process of discovering knowledge within data.

Topics Covered:
- What is data mining?
- What is data warehousing?
- Data mining functionalities - what kinds of patterns can be mined?
- Data preprocessing, cleaning, integration, reduction and transformation, data reduction
- Discretization and concept hierarchy generation
- Data mining primitives, languages, and system architectures
- Attribute analysis
- Association rule mining
- Classification by decision tree induction
- Clustering
- Issues in contemporary data mining applications.
Student Learning Outcomes:
Upon successful completion of this course, the student will be able to:

- Identify data mining functionalities
- Identify data warehousing functionalities
- Apply data preprocessing techniques - data cleaning, data integration and transformation, data reduction, discretization, and concept hierarchy generation
- Describe data mining primitives, languages, and system architectures
- Use a Data Mining Query language
- Apply analytical characterization techniques
- Perform attribute relevance analysis
- Mine descriptive statistical measures in large databases
- Mine association rules in large databases
- Use classification by decision tree induction techniques on large databases
- Use data mining software to perform data mining functionalities – mining association rules, classification and prediction and decision tree analysis.
- Identify current social issues in data mining
- Describe the current needs in data mining research

Required Textbook(s):
Data Mining: Concepts and Techniques, Third Edition by Jiawei Han, Micheline Kamber and Jian Pei, 2012. ISBN: 978-0-12-381479-1.

References:
Weka’s site: http://www.cs.waikato.ac.nz/~ml/weka/

Tentative grade distribution:

<table>
<thead>
<tr>
<th>Component</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Quizzes</td>
<td>15%</td>
</tr>
<tr>
<td>Midterm</td>
<td>15%</td>
</tr>
<tr>
<td>Final exam</td>
<td>20%</td>
</tr>
<tr>
<td>Assignments</td>
<td>20%</td>
</tr>
<tr>
<td>Project</td>
<td>30%</td>
</tr>
</tbody>
</table>

Technology Requirements:
Knowledge of a machine learning tool – WEKA (on the Windows environment) will be necessary for the project.

Expectations for Academic Conduct/Plagiarism Policy:
Academic Conduct Policy: [Web Site]
Plagiarism Policy: [Web Site]
Student Handbook: [Web Site]

Assistance:
Students with special needs who require specific examination-related or other course-related accommodations should contact the Director of Disabled Student Services (DSS), dss@uwf.edu, (850) 474-2387. DSS will provide the student with a letter for the instructor that will specify any recommended accommodations.

Exams: Makeup exams will NOT be given except with a serious, documented medical or legal excuse. No makeups will be given unless students make advance arrangements. The final exam is comprehensive.
Re-grading Assignments: It is the student’s responsibility to check graded assignments/tests when they are returned to you. I will gladly re-grade an assignment/test when a question or mistake is brought to my attention. To ensure fairness, I reserve the right to re-grade the entire assignment/test. As a result, your grade may increase, decrease, or remain the same. Grades will not be changed after a week from the date graded assignments/tests are returned to the class.

Grades: Final grades will be calculated using a standard grade distribution.

- **March 21st**: The last day of the term for withdrawal from an individual course with an automatic grade of “W”.
- **April 25th**: Deadline for students requesting late withdrawal (W or WF at my discretion) from class, must have the approval of the advisor, instructor, and department chairperson (in that order) and finally, by the Academic Appeals committee. Requests for late withdraws may be approved only for the following reasons (which must be documented):
  1. A death in the immediate family.
  2. Serious illness of the student or an immediate family member.
  3. A situation deemed similar to categories 1 and 2 by all in the approval process.
  4. Withdrawal due to Military Service (*Florida Statute 1004.07*)
  5. National Guard Troops Ordered into Active Service (*Florida Statute 250.482*)

- Requests without documentation will not be accepted. Requests for late withdrawal simply for not succeeding in a course do not meet the criteria for approval and will not be approved.

- Applying for an incomplete or “I” grade will be considered only if: (1) there are extenuating circumstances to warrant it, AND (2) you have a passing grade and have completed at least 70% of the course work, AND (3) approval of the department chair.

Participation and Feedback: I encourage active participation in class. I believe that effective communication between the instructor and students will make the course more useful, interesting, and productive. **Please contact me if you have any questions, concerns, or suggestions! 😊**

Social Media Policy: In order to provide fairness and appropriate professional separation between instructor and students, I will not accept social media invitations from students.

Important Note: Any changes to the syllabus or schedule made during the semester take precedence over this version. Check the eLearning site (or email) regularly for up-to-date information.
Overall Grading Scale:
A : 89.5 – 100
B+ : 86.5 – 89.499
B : 79.5 – 86.499
C+: 76.5 – 79.499
C : 69.5 – 76.499
D+: 66.5 – 69.499
D : 59.5 – 66.499
F : 0 – 59.499

Late Policy:
1. You are expected to complete work on schedule. Deadlines are part of the real world environment you are being prepared for.
2. Documentation of health or family problems may be required.
3. If you have to miss a class, be sure you arrange with another student to find out what you missed at the earliest possible date.
4. Late assignments will not be accepted.

There’s another page...keep scrolling down...
Tentative Course Schedule:

<table>
<thead>
<tr>
<th>WEEK #</th>
<th>DATE</th>
<th>TOPIC</th>
<th>PROJECT</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Jan 6</td>
<td><strong>Chapter 1</strong> – Introduction, What is data mining?</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>Jan 13</td>
<td><strong>Chapter 2</strong> – Getting to know your data</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>Jan 20</td>
<td><strong>Chapter 3</strong> – Data Preprocessing</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>Jan 27</td>
<td><strong>Chapter 4</strong> – Data Warehousing and Online Analytical Processing</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>Feb 3</td>
<td><strong>Chapter 6</strong> – Mining Frequent Patterns, Associations, and Correlations</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>Feb 10</td>
<td><strong>Chapter 8</strong> – Classification: Basic Concepts</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>Feb 17</td>
<td><strong>Chapter 9</strong> – Classification: Advanced Methods</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>Feb 24</td>
<td><strong>Midterm Exam</strong></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>Mar 3</td>
<td><strong>Chapter 10</strong> – Cluster Analysis</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>Mar 17</td>
<td><strong>Chapter 12</strong> – Outlier detection</td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>Mar 24</td>
<td><strong>Chapter 13</strong> – Data mining trends and research frontiers</td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>Mar 31</td>
<td>Data Mining Applications</td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>Apr 7</td>
<td>Project Q and A, time to work on project</td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>Apr 14</td>
<td>Project Q and A, time to work on project</td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>Apr 21</td>
<td>Review for final exam</td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>Apr 28</td>
<td><strong>Final Exam, Comprehensive</strong></td>
<td></td>
</tr>
</tbody>
</table>

**Important Note:** Assignments will be assigned throughout the semester. Due dates for assignments will be given in class.

Also, there is a major project in this class. The project will be a semester project that will have many parts. As the semester progresses, I will give you milestones when each part of the project will be due. Students are expected to do all parts of the project and **all parts of the project will receive a grade** (adding to up 30% in the end).

Also, tests and quizzes will be announced as the semester progresses.

*Enjoy the semester...*

*Dr Snider*