CSC 466 - Knowledge Discovery from Data

1. CSC 466 Knowledge Discovery from Data

2. 
   
   credit units 4  
   contact hours 6

3. Course Coordinator: Lubomir Stanchev

4. 
   Textbook (or other required material): Bing Liu, Web Data Mining or Zaki Mohammed and Meira Wagner, Data Mining and Analysis

5. 
   a. Course Description: Overview of modern knowledge discovery from data (KDD) methods and technologies. Topics in data mining (association rules mining, classification, clustering), information retrieval, web mining. Emphasis on use of KDD techniques in modern software applications.
   b. Prerequisite: CSC 349 and one of the following: STAT 302, STAT 312, STAT 321 or STAT 350.
   c. Required/Elective/Selective Elective for CPE, CSC, SE

<table>
<thead>
<tr>
<th></th>
<th>CSC</th>
<th>CPE</th>
<th>SE</th>
<th>CDSM/DATA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Required</td>
<td></td>
<td></td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>Elective</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>Selective Elective</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

6. 
   a. Course Learning Objectives

   The student will be able to:
   • Implement and apply algorithms for documents search using natural-language queries.
   • Evaluate the quality of a document search algorithm.
   • Extract association rules from a data set.
   • Implement and apply Supervised Learning algorithms.
   • Implement and apply Unsupervised Learning algorithms.
   • Evaluate Supervised and Unsupervised learning algorithms.
   • Implement and apply different Link Analyses algorithms.
   • Implement and apply Collaborative Filtering algorithms.

   b. Level at which Student Outcomes are addressed
   (“B” = Basic level, “I” = Intermediate level, “A” = Advanced level)

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
</tr>
</thead>
<tbody>
<tr>
<td>CSC</td>
<td>A</td>
<td>A</td>
<td></td>
<td></td>
<td>I</td>
<td>A</td>
<td></td>
</tr>
<tr>
<td>SE/CPE</td>
<td>A</td>
<td>B</td>
<td></td>
<td></td>
<td>I</td>
<td>A</td>
<td>B</td>
</tr>
</tbody>
</table>

7. Major Topics Covered: (number of lecture hours per)
• Introduction to Information Retrieval. (6)
• Link Analyses (3)
• Association Rules (3)
• Supervised Learning (9)
• Unsupervised Learning (6)
• Collaborative Filtering (3)