CSC 566 – Topics in Advanced Data Mining

1. CSC 566 – Topics in Advanced Data Mining

2. **credit units**: 4  **contact hours**: 4

3. **Course Coordinator**: Alex Dekhtyar

4. **Textbook (or other required material)**: None

5. a. **Course Description**:
   Advanced topics in the areas of data mining, knowledge discovery in data, machine learning, information retrieval and intelligent analysis of information. The Class Schedule will list topic selected. Total credit limited to 8 units. 4 lectures.

   b. **Prerequisite**: CSC 466 or CSC 480 or CSC 582.

   c. **Required/Elective/Selective Elective for CPE, CSC, EE, SE**

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

6. a. **Course Learning Objectives**
   The student will be able to:
   - Implement advanced knowledge discovery in data methods, techniques, algorithms to analyze data.
   - Analyze data using advanced knowledge in data algorithms, methods, and techniques.
   - Describe the work of advanced knowledge in data methods, techniques, and algorithms.
   - Understand societal impacts of knowledge discovery in data, and apply this understanding to inform technical decisions.
   - Design and perform empirical studies for establishing the accuracy and performance of knowledge discovery in data methods, techniques, and algorithms and their implementations.

   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>A</td>
<td>A</td>
<td>I</td>
<td>A</td>
<td>N/A</td>
</tr>
<tr>
<td>SE/CPE</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>I</td>
<td>A</td>
<td></td>
</tr>
</tbody>
</table>
7. **Major Topics Covered: (number of lecture hours per)**
   - Classification problem (2 hours)
   - Support Vector Machines (3 hours)
   - Dimensionality reduction techniques in machine learning (3 hours)
   - Principal Component Analysis and its uses in KDD (2 hours)
   - Latent Semantic Indexing (1 hour)
   - Generative Probabilistic Methods (8 hours)
   - Artificial Neural Networks (2 hours)
   - Clustering (5 hours)