CSC 482 – Speech and Language Processing

1. CSC 482 – Speech and Language Processing

2. credit units  4    contact hours  6

3. Course Coordinator: Foaad Khosmood

4. Textbook (or other required material):
   Jurafsky and Martin, "Speech and Language Processing".
   Bird, Klein, Loper, "Natural Language Processing with Python".

5. a. Course Description:
   Introduction to natural language processing theory; speech processing; review of recent
   advancements. Topics include: tokenization, part-of-speech tagging, word-sense
   disambiguation, natural language understanding, natural language generation, data mining,
   voice processing, vocalization, semantic networks, intelligent assistants, computational
   linguistics, stylistics and machine learning. 3 lectures, 1 laboratory.

   b. Prerequisite: **CSC 466** or **CSC 480** or graduate standing. Recommended: **CSC 349**.

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

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

6. a. Course Learning Objectives
   The student will be able to:
   • Create natural language statistical models.
   • Find, locate, and create Natural Language Processing (NLP) corpora for analysis.
   • Intake and perform preprocessing pipeline operations on NLP corpora.
   • Apply computational methods to perform statistical parsing.
   • Apply computational lexical semantics.
   • Evaluate and implement systems to address contemporary NLP challenges such as
     semantic networks, machine translation, question answering systems, vocalization,
     and automatic summarization.
   • Explain contemporary theory and practice of voice processing.
   • Apply machine learning and deep learning methods to NLP.

   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>1</td>
<td>N/A</td>
<td></td>
</tr>
</tbody>
</table>
7. **Major Topics Covered: (number of lecture hours per)**
   - N-grams and preprocessing (3 hours)
   - Part-of-Speech Tagging (3 hours)
   - Context Free Grammars and Parsing (3 hours)
   - Statistical Parsing (3 hours)
   - Lexical Semantics and Computational Methods (3 hours)
   - Question Answering Systems and Summarization (2 hours)
   - Text Classification by Machine Learning (3 hours)
   - Machine Translation (1 hour)
   - Speech Processing (1 hour)
   - Deep Learning Methods (1 hour)