1. **Catalog Description of the Course.**

   **COMP 578 DATA MINING (3)**
   Three hours of lecture in the lab per week.
   **Prerequisite:** Admission to the Computer Science or Mathematics Graduate Program

   This graduate course covers the fundamentals of Data Mining. Topics include: the analysis of patterns of data in large databases and data warehouses, the application of statistical pattern recognition, and data modeling and knowledge representation. Applications in large databases and gene hunting.

2. **Mode of Instruction.**

<table>
<thead>
<tr>
<th>Units</th>
<th>Hours per Unit</th>
<th>Benchmark Enrollment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lecture</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>Seminar *</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>Laboratory</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>Activity *</td>
<td>*</td>
<td>*</td>
</tr>
</tbody>
</table>

3. **Justification and Learning Objectives.**

   **Justification:** Elective in the MSCS program.
   **Learning Objectives:**
   
   Through this course, students will be able to:
   
   1. Index large data sets
   2. Design search engine for text databases
   3. Design search engine for multimedia
   4. Build probabilistic models for the knowledge content of data
   5. Apply data mining techniques to gene hunting and proteomics
   6. Use BLAST programs comprehensively.

4. **Is this a General Education Course?**
   No.

5. **Course Content in Outline Form.**

   **Topics:**
   a) the analysis of patterns of data in large databases and data warehouses
   b) deterministic and probabilistic techniques for constructing hashkey for large data sets
   c) the application of statistical pattern recognition
   d) data modeling and knowledge representation
   e) Applications in large databases, gene hunting.

6. **References.**

   **Texts:**
   *Data Mining with Applications and Implementation in Java*, I. Witten, D. Frank, Morgan Kaufman (2002)

7. **Faculty Qualified to Teach This Course.**

   Qualified Faculty: CS Faculty

8. **Frequency.**

   Projected semesters to be offered: Spring X

9. **New Resources Required.**
a. New Equipment needs: Use of the existing labs
b. New Library needs: None
c. New Space/Facilities needs: None

10. Consultation.
    Attach consultation sheet from all program areas, Library, and others (if necessary).

11. Program Modification.
    If this new course will alter any degree, credential, certificate, or minor in your program, attach a
    program modification.

12. Proposer of Course.
    This course is an elective for graduate students in MS in Mathematics and MS in Computer
    Science programs.

Proposer  Date

10/31/2003