CS 697 Spatial Data Mining

Exam Review
Spring 2009

- **Schedule**
The examination is open-book and open-note.

There are **120 minutes** for the Examination.

<table>
<thead>
<tr>
<th>Exam Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>7:10 PM – 9:10 PM Wednesday April 29</td>
</tr>
</tbody>
</table>

- **Preparation Materials**
Use lecture notes as your major preparation materials. Students are strongly recommended to go through all the examples used by the lecture notes.

Note: If there are any inconsistency between the lecture notes and the textbook, use class lecture notes.

- **Topics**
1. (20%) Data Preprocessing
   - Type of attributes
   - Different techniques of data preprocessing
2. (50%) Classification
   - Definition of classification
   - Hunt’s algorithm
   - How to split the records during tree induction
   - Gini index
   - Entropy
   - Misclassification error
   - Underfitting and overfitting
   - How to address overfitting
   - Metrics for performance evaluation
   - Methods for model comparison
3. (30%) Association analysis part I
   - Definition of association rule
   - Frequent itemset generation using Apriori Principle
   - Maximal frequent itemsets
   - Closed frequent itemsets