COSC 4335 *“Data Mining”*

Review Sheet for 2015 Final Exam

The final exam has been scheduled for Tu., May 10, 2p in our class room and will be “open everything” and will take about 100 minutes. The use of computers is not allowed! The exam counts approx. 20% towards the final course grade. The exam will cover the following topics:

1. \*\*\*\*\* Classification Techniques
   1. all transparencies covered in the lecture which cover decision trees, overfitting, support vector machines, ensembles and n-fold cross-validation.
   2. book pages: 147-164 168-183 186-187, 256-274, 276-280, 285-290
2. \*\*\*\* Association Analysis
   1. book pages: 327-341, 349-353, and all transparencies that are associated with those book pages (Part2 of the association analysis transparencies is **not** relevant for the final exam)
3. \*\* PageRank
   1. Class transparencies excluding *Gleich Dissertation transparencies*
   2. Wikipedia PageRank page

Comment: will only ask very basic questions about this topic.

1. \*\*\*\* Outlier Detection
   1. Transparencies covered in the lecture
   2. Textbook pages 651-662; 665-667 (subsection 10.2.3 is **not** relevant for the final exam!)
2. \*\*\* Preprocessing
   1. Preprocessing transparencies
   2. Textbook pages 36-59

Algorithms and techniques you should know well include: SVM model generation, decision tree induction, Adaboost algorithm, APRIORI algorithm for association rule mining, k-nearest neighbor based outlier detection.