Computer Engineering Department
Bilkent University

CS533
Fall 2018
Midterm Exam Study Guide
November 15, 2018

A. Exam Structure - Time Schedule & Other Information

1. Mostly problem solving type questions. No. of questions: 7-8 (approx.).

2. Closed everything exam.

3. Exam may involve simple calculations. You may bring a simple calculator. You cannot share calculators and are not allowed to use your cellular phones.

4. Exam Time (Thursday, November 22, 2018) 15:40 – 17:15

5. Note that we have a mock midterm exam available on our course web site.

6. Don't miss the class notes on our course web site.

B. Subjects to be included (an <u>incomplete</u> list)

1. No need to memorize the formulas (most of them will be provided: CC and similarity formulas; you have to know recall and precision by hard).

2. Similarity calculations for inverted files.

3. Inverted file concepts incluZobel-Moffat paper from *ACM Computing Surveys*: Inverted Indexes for Search Engines.

4. Evaluation of IR systems.

5. Clustering concepts.

6. Clustering algorithms: heuristic … complete link.

7. Cover coefficient based methodology, C^3M.

8. Cluster validation: Yao's formula, Rand index.

9. Maximal marginal relevance.

10. Data fusion: Reciprocal rank, Borda Count, Condorcet.

11. Stemming.

12. Please also consider the papers that you have read for solving the assignment no. 1 and 2 questions. About 15 points for this part.

    Subjects beginning with term weighting topic and beyond are not included, like Salton-Buckley classic term weighting paper, TDV concept, PAT tress, Signature Files.