

**Computer Engineering Department  
Bilkent University**

**CS533  
Fall 2019  
Midterm Exam Study Guide  
November 7, 2019**

**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 (Tuesday, November 12, 2019) 10:40 – 12:15
5. Don't miss the class notes on our course web site.

**B. Subjects to be included (an incomplete list)**

1. No need to memorize the formulas (most of them will be provided: Cover coefficient and similarity formulas; you have to know recall and precision by hard).
2. Similarity calculations for inverted files.
3. Inverted file concepts including Zobel-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<sup>3</sup>M.
8. Cluster validation: Yao's formula, Rand index.
9. Data fusion: Reciprocal rank, Borda Count, Condorcet (my class presentation was intentionally incomplete, please study it to its full extend).
10. PAT trees.
11. Please also consider the paper A. K. Jain, M. N. Murty, P. J. Flynn, "Data Clustering: A Review" that you have read for assignment no. 1. Skim the paper about 10 points for this part.