Introduction to
Information Retrieval

CS276
Information Retrieval and Web Search
Christopher Manning and Pandu Nayak
Lecture 1: Introduction

Course logistics in brief

- Instructors: Christopher Manning and Pandu Nayak
- TAs:
  - Ashwin Paranjape (Head TA)
  - Chris Chute
  - Fei Jia
  - Rohan Sampath
- Time: TuTh 4:30–5:50, Gates B01 ($\rightarrow$ SCPD video)
- Class webpage: http://cs276.stanford.edu/
  - Will have office hours, etc. (starting next week)

2

Work for the class

- Required work:
  - 2 problem sets @ 10% = 20%
  - 3 programming assignments, 3rd one larger than first two:
    - @ 13%, 13%, 20% = 46%
  - Final exam: 30%
  - Class participation: 4%
    - 2% for (on-campus) attending guest lectures in person or (SCPD, unavoidable absences) writing reaction paragraph
    - 2% for Piazza participation, mid-quarter survey completion, and/or being present and active in class
- Problem sets and assignments
  - Due at: 4pm, either Tue (early in course) or Thu (later on)
  - Assignments in Python this year

3

What do we hope to teach?

- How to do efficient (fast, compact) text indexing
- Retrieval models: Boolean, vector-space, probabilistic, and machine learning models
- Evaluation and IR interface issues
- Document clustering and classification
- Search on the web, including crawling, link-based algorithms, indirect feedback, metadata, and personalization

4