Algorithms for Modern Data Sets
(COMP 3801)

Anil Maheshwari
anil@scs.carleton.ca
School of Computer Science
Carleton University
Main Topics

Required Background

HyFlex Teaching

Course Evaluation

Help Me
Main Topics
<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>How search engines search?</td>
</tr>
<tr>
<td>2.</td>
<td>How do recommendation systems recommend?</td>
</tr>
<tr>
<td>3.</td>
<td>How to find a matching fingerprint?</td>
</tr>
<tr>
<td>4.</td>
<td>How to find similar documents?</td>
</tr>
<tr>
<td>5.</td>
<td>How to detect junk emails?</td>
</tr>
<tr>
<td>6.</td>
<td>How to find what is trending in last week or last month?</td>
</tr>
<tr>
<td>7.</td>
<td>How to find nearest neighbor of a query point in high dimensional data?</td>
</tr>
<tr>
<td>8.</td>
<td>How to find communities on the web?</td>
</tr>
<tr>
<td>9.</td>
<td>. . .</td>
</tr>
</tbody>
</table>
Our Focus

- Algorithmic Techniques
- Key Ideas
- Learn bit of Math
- High-level details
Required Background
Background

- **Basic Data Structures**
  Lists, Stacks, BST, Hashing, Searching and Sorting
  See Pat Morin's [https://opendatastructures.org/](https://opendatastructures.org/)

- **Basic Probability Theory**
  Expectation, Variance, Concentration Bounds
  Joe Blitzstein [https://projects.iq.harvard.edu/stat110/youtube](https://projects.iq.harvard.edu/stat110/youtube)

- **Linear Algebra**
  Eigenvalues and Eigenvectors, Null Space, Positive Definitive Matrices, $Q \Lambda Q^T$, SVD
  Gilbert Strang

- **Algorithmic Techniques**
HyFlex Teaching
(Off/On)line Course

- Lectures twice a week over zoom either from the HyFlex room or from my office
- Ample references will be provided
  2. Notes on Algorithm Design on my web-page
     [http://people.scs.carleton.ca/~maheshwa/](http://people.scs.carleton.ca/~maheshwa/)
  3. Numerous research articles
     (accessible from Carleton Library/Web).
  4. Pointers to useful Videos.
Office Hours

Likely during and after the class, and setup over e-mail.
Course Evaluation
## Evaluation Parameters

### Principles
- Open-ended course by design
- Contents evolve
- Enjoyable learning experience
- Use forums effectively

### Evaluation Components
- 4 Assignments: 40%
- 2 Tests: 24%
- Final Exam: 36%
Help Me
1. Post your questions on Forums
2. Post useful links/references/ideas
3. Take initiative to learn
4. E-mail now if you want to learn some specific algorithmic technique
5. Seek Help, Have Fun & Smile