### Topics

+ reviewing string clustering

+ writing

+ spatial clustering

+ scalabled retrieval / recommendation

# --------------- writing

### Plagiarism / Citation

+ Writing with citation / source / reference is good practice

+ Copying from websites or papers without citation is considered as stealing / plagiarism (lack of creativity).

+ Select good / trusted sources to gain more confidence in your writing

### Structure of a technical report / thesis / research paper

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| introduction  (Tổng quan) | literature review  (Các phương pháp có) | your method  (Phương pháp mới của bạn) | conclusion | references  citations |
| + 01 chapter  + 1 – 2 pages | + 01 chapter  + 5 – 10 pages | + 01-03 chapters  + 10 – 40 pages | + 01 chapter  + 0.5 to 2 pages | + 01 chapter  + no limit |
|  | **existing works**  **past works** | **current work** | **fututre work** |  |
| + what is your problem?  >> recommendation  >> applied recommendation  + what is wrong with existing methods? what are limitations with existing methods? **Why do they fail?**  >> youtube is good?  >> amazon recommendation is good enough?  >> code is accessible?  **+ why does your method work?** What is the key to your success? >> foundation of your method >> imdb uses user information >> we use meta data  + overview of your method and framework | + baseline method (closely-related method)  >> kmeans (baseline) >> kmeans for movies kmeans + movsimf kmeans + movsimf + adjmean  >> kmeans for strings   advanced method  = baseline  + something new  + development trend BOW (199?) pLSA (1999) LDA (2001)  + read the newest paper / book and trace back | ------ describe  + describe your method  + can they implement?  ------ theoretical analysis  + big-oh time  + convergence  + why faster?  + why better?  ------ test  + describe your testing  (experimental results)  + empirical evidence | + what is the problem again?  +what have been done?  >> summarize what you have done  + what is next?  >> future of your work |  |

### Some case

+ poisson editing (2003) >> photoshop did it before (2001)

+ green-screen editing (1995) invented, first paper (1997)

### Using words

# ---------------- spatial classification / clustering

### Spatial kNN classification

|  |  |
| --- | --- |
|  | + You are given the image with labeled points  + You have to label the rest based on the example points.  + Choose a method, you can try kNN classification |

### Spatial kmeans clustering

|  |  |
| --- | --- |
| C:\Users\Administrator\Desktop\auto-divide-4.png | + you are given points in space  + you need to divide them into clusters into 4 groups |