**Kiến trúc hệ thống:**

* + PaperRecommender
    - Source package
    - Constant
      * Các hằng số
    - Document
      * Tài liệu
    - uit.tkorg.paperrecommender
      * GUI
      * Controller
        + Central

Controller trung tâm: trực tiếp nhận yêu cầu từ GUI, điều phối logic chương trình, gọi các sub module phù hợp trong từng component.

Main entry dùng khi develop, test.

* + - * + Recommendation

Logic to recommend.

* + - * + Evaluation

Logic to evaluate.

* + - * Model
        + Represent data.
      * Utility
        + Class tính toán, logging, input, output, ...

**Module:**

Các module chính:

* Data utility
  + Generate vocabulary, save to file (serialize), load from file.
* Compute papers' feature vector.
  + Generate list of papers.
  + For each paper (class paper -> object)
    - Read paper content
    - Generate basic feature vector (paper content only).
  + Compute weight:
    - Cosine
    - *RPY*
  + Compute papers' full feature vector (3 type: linear, cosine, rpy)
* Compute researchers' profile vector.
  + Generate list of authors, each type and the whole, save, load.
  + For each author (class author -> object)
    - Data: list of paper, list of paper citation, ref.
    - Method:
      * Arrange list of papers
      * Compute profile vector by combining papers' feature vector.
        + For each weighting scheme.
  + Junior class inherits author
  + Senior class inherits author
* Generate recommendation list:
  + For each combination of feature vectors:
    - Compute similarity of researchers and papers.
    - Generate ranked list of recommend papers.
    - Save generated list to file.
* Compute evaluation metric:
  + Load ground truth list.
  + Load recommendation list.
  + Integrate library to compute evaluation metric.
  + Compute NDCG@5, 10; MRR.

**Note:**

* Each data class has to implement Serializable interface.
* Project architecture:
  + Using ordinary java project in netbeans (java application, ant build).
  + Not using maven. So not using module style.
  + Using MVC 2:
    - Each package m, v, c could contain many package to group class by function.

**Framework:**

**1.**

**DATA IMPORT**

**2.**

**DATA PREPARATION**

**3. RECOMMENDATION**

**4. EVALUATION**

Dataset 1 Import

Author and paper objects

DS1

Content-based

NDCG

MRR

Content-based

Author and paper feature vector

File content

Utility

Weighting

Evaluation

Recommendation rank list

Compute weighting

Compute evaluation metric

Ground-truth list

DS

Evaluation score

Dataset 2 Import

DBLP Import

CF

CF

MAP