

## Một số hướng dẫn về nội dung Seminar môn KTDL

(ver. 2016/10/24)

### 1. Hình thức thực hiện (nhóm/cá nhân):

Thực hiện theo nhóm, mỗi nhóm 3 thành viên.

### 2. Nội dung công việc:

- a. Tìm hiểu một bài báo khoa học (scientific paper) tiếng Anh.
- b. Viết tóm tắt cách hiểu của bản thân về bài báo ở dạng một tóm tắt khoa học (extended abstract).
- c. Soạn và trình bày slides trước lớp. SV phải hiểu rõ bài báo đã chọn để giải thích cho các sv khác hiểu và đồng thời có thể trả lời các câu hỏi.

### 3. Các yêu cầu cơ bản:

- a. Các bài báo được chọn phải là bài được công bố tại hội nghị/tạp chí có uy tín hoặc được trích dẫn bởi nhiều bài báo khác. Bài báo phải là bài báo dài (long paper). Gợi ý một số hội nghị tốt mà SV có thể chọn bài: AAI, IJCAI, KDD, WWW, WSDM, ICDM, ICDE, CIKM, ISWC, EKAU, KSEM, DASFAA, PAKDD, SDM, DaWak, PKDD.
- b. SV có thể chọn bài báo tại các tạp chí khoa học uy tín, tuy nhiên bài báo tại hội nghị được khuyến khích hơn.
- c. Tóm tắt phải từ 1000 đến 1500 chữ. Tóm tắt bao gồm đặt vấn đề (bài toán là gì, đâu là điểm cần giải quyết), cách giải quyết có gì đặc biệt (ưu điểm của cách giải quyết), tóm lược cách giải quyết, cách đánh giá, kết quả đánh giá.
- d. Sinh viên phải làm việc theo nhóm, có sự phân công rõ ràng, có lịch trình làm việc hợp lý, mỗi thành viên phải đảm bảo hoàn thành công việc được giao. Dựa trên lịch làm việc, các báo cáo, công việc của mỗi thành viên, ghi chú họp nhóm đều phải công khai cho giáo viên theo dõi và theo đúng tiến độ đã lên lịch.
- e. SV được khuyến khích hiểu rõ bài báo và các kiến thức nền sử dụng trong bài báo. Tuy nhiên, các kiến thức nền hoặc các vấn đề liên quan có thể rất rộng lớn. Tùy theo mức độ ảnh hưởng của kiến thức nền đối với bài báo mà SV cần tìm hiểu chi tiết/tổng quát về nó chứ không nên bỏ qua (chỉ xem các bước/thành phần là hộp đen khi nó không quan trọng hoặc quá phức tạp).
- f. Trong quá trình thực hiện, sinh viên có thể gặp trực tiếp GV khi cần hỗ trợ. SV liên hệ qua email để đặt lịch hẹn và ghi rõ thời gian mong muốn. Thời gian: 8h đến 11h, T2 đến T5, phòng I81 (bộ môn Khoa học máy tính). SV cần giúp đỡ của GV để lấy bài báo có thể gửi mail cho GV.  
Địa chỉ email : [nhkhai@fit.hcmus.edu.vn](mailto:nhkhai@fit.hcmus.edu.vn)

#### 4. Các mốc thời gian (dự kiến):

- a. 3/11: SV nộp 3 bài báo theo thứ tự ưu tiên (tác giả, tựa đề, hội nghị, abstract) để GV chọn ra 1.
- b. 5/11: Phản hồi của GV cho từng nhóm. SV nộp tiếp 1 bài báo khác nếu GV không chọn được bài nào phù hợp.
- c. 9/11: SV nộp phân công và lịch làm việc cơ bản (lịch làm việc/phân công có thể thay đổi trong suốt quá trình). GV công bố các chủ đề đã chọn.
- d. 25/11: SV nộp báo cáo (tóm tắt), GV công bố cho cả lớp.
- e. 2/12: SV nộp slides (slides thuyết trình chính thức có thể điều chỉnh khi cần thiết (nên hạn chế)).
- f. Từ 9/12: Seminar tại lớp.
- g. 9/1: Nộp slides hoàn chỉnh.

#### 5. Hình thức và yêu cầu về seminar:

- a. Giáo viên công bố lịch trình bày vào đầu mỗi buổi (các nhóm chưa thuyết trình đều phải trong tư thế sẵn sàng).
- b. Thời gian cho mỗi nhóm là 30 phút (25 phút trình bày, 5 phút hỏi đáp). SV được yêu cầu phải luyện tập từ trước và canh chỉnh đúng thời gian được cho phép (không dài hơn, không ngắn hơn).
- c. Các bài trình bày phải vừa đúng 25 phút.
- d. Khuyến khích tất cả thành viên đều trình bày (ví dụ mỗi sv 5-10 phút).
- e. SV khán giả phải chú ý theo dõi các bài trình bày của nhóm khác để tiếp thu kiến thức và đặt câu hỏi. Trường hợp không có nhóm nào đặt câu hỏi, GV có quyền chỉ định bất kỳ các nhóm đặt câu hỏi.
- f. Nhiều bài báo có slides trình bày và video trình bày tại hội nghị (videolecture), SV được khuyến khích tham khảo các tài liệu đó, nhưng không được phép tái sử dụng slides.

#### 6. Đánh giá:

- a. Đánh giá tiến độ làm việc của nhóm và từng thành viên thông qua lịch làm việc, phân công, tuân thủ các mốc thời gian.
- b. Nội dung báo cáo (tóm tắt).
- c. Nội dung slides.
- d. Bài trình bày tại lớp và trả lời câu hỏi.
- e. Tham gia thảo luận, đặt câu hỏi khi các nhóm khác trình bày.

#### 7. Gọi ý hiểu bài báo

- a. Hiểu tầm quan trọng và các ứng dụng của chủ đề đã chọn, ngay cả khi ứng dụng chưa được hiện thực hóa ở thời điểm hiện tại. Các dữ liệu nào được quan tâm nhiều trong bài toán này (raw text, relational data, graph, image, video, music, metadata, social media, linked data, ontology...).

- b. Hiểu các mô hình của chủ đề đã chọn. Ví dụ, đối với data integration có ba mô hình thông dụng là dựa trên máy học (classification-based), dựa trên đặc tả so khớp (link specification), và mô hình lặp (iterative).
- c. Tùy từng mô hình mà các bài toán nhỏ là khác nhau. Cần hiểu các bài toán nhỏ (tương ứng với các bước) của mô hình đã chọn ở bước 2, và xác định vấn đề nào đang được quan tâm. Chẳng hạn như phát triển framework, thuật toán, hệ thống, phương pháp, kỹ thuật...
- d. Bài báo có thể giải quyết một bài toán nhỏ hoặc đề xuất một hệ thống/mô hình cho bài toán lớn hơn. SV cần nắm rõ bài toán mà bài báo quan tâm là gì, lý luận nào cho việc đề xuất một bài báo như vậy (đâu là đóng góp khoa học/áp dụng thực tiễn của bài báo). Thực nghiệm trên dữ liệu gì, kết quả và đánh giá kết quả.