

Assignment 2: Independent Reading Study

Due date: Tuesday, March 1

Note: Additional due dates apply, see full description of deliverables and due dates below.

Assignment

This is an **team assignment**.

I have selected six topics for independent study in this course, formed six three-to-four student teams and assigned each team an independent study topic. For the sake of completeness, the list of assigned topics is:

- Hidden Markov Models and their applications to Machine Learning and Data Mining
- Analysis of Data Streams
- Data Mining/Machine Learning for Security Applications
- Advanced Pattern/Association Mining
- Recommendation Systems
- Generative Probabilistic Models and Topic Modeling

Over the course of the quarter, each team will perform the following activities.

- **Select the readings.** Each team (possibly, in collaboration with the instructor) will select a set of readings for the assigned topic.
- **Study the readings.** Self-explanatory. Read the materials, discuss them in the group.
- **Prepare Lecture Notes.** Prepare a set of succinct but comprehensive lecture notes to cover and introduction into the topic of your study.
- **Prepare Lecture.** Prepare a 30-35 lecture on the topic of your study, and deliver it.

Deliverables and Deadlines

Your activities have the following deliverables and deadlines. The short summary is in the tables below. Detailed instructions on each deliverable follow.

Deliverable	Format	Due Date
Proposed list of readings	Googledocs	January 27, Thursday
Mid-Quarter Check-in	PDF	February 15, Tuesday
Lecture Notes (final draft)	LaTex and PDF	March 1, Tuesday
Lecture	Oral presentation	March 8 (Tuesday), March 10 (Thursday)
Lecture Notes (final version)	LaTex and PDF	March 15 (Tuesday)

In addition to tangible deliverables, each team will have the following interactions with me:

Topic	Timing
Comments/Approval of proposed list of readings	January 31 (Monday) - February 3 (Thursday)
Progress report	February 14 (Monday) – February 17 (Thursday)
Lecture preparation	February 28 (Monday) – March 3 (Thursday)

Assignment Stages

Your assignment contains five tangible deliverables and at least three scheduled interactions with the instructor.

Proposed List of Readings

Due Date: January 27, Thursday.

On the first stage of the assignment each team needs to prepare a reading list. The reading list can consist of material from any available to the team textbooks (or books in general), or academic papers. Some on-line materials and videos can be used as *support* for the academic reading materials (papers and books), but these materials cannot be the only things on your list¹

The list shall be organized as a collection of bibliographic references with links to any on-line resources (including PDFs of papers) where available. The final version of the reading list will eventually turn into a LaTeX bibliography (so you can start maintaining your BibTeX files now), but for the purpose of the first milestone, create a GoogleDocs document in which all included URLs are clickable. In addition, if you already collected all the proposed reading materials, create a Google Drive directory, and place all materials in the directory, and share the directory with me.

Approval process. I will review the proposed readings over the January 29-30 weekend, and will have prepared comments for each team starting Monday, January 31. Some comments may be shared pre-emptively via Slack, but in addition (and unless I explicitly indicate to a specific team that I do not need to meeting with it) I would like to meet with each team for 10-20 minutes to discuss the proposed reading list and the agenda for the lecture the team will be preparing. The meetings can take place during any scheduled office hours (Monday 9-10, Tuesday, 8-9, Wednesday, 8-10), or at any other agreed upon time - if you cannot make it to the office hours, contact me via Slack and we will set up a meeting time.

The results of our conversation may entail a change of plans/updates to the reading lists. If such updates are significant and need some off-line work for a team (e.g., an effort to find an additional paper on a specific topic), the team will have a three day period from the time of the meeting to complete the reading list revisions.

Mid-Quarter Check-in.

Due Date: February 15.

Between the end of the approval process and February 15 is your time to read the reading materials from your reading lists, and to discuss general plans for your lecture. I recommend that each team has a team meeting prior to February 15 to lay out the plans for the lecture. The mid-quarter check-in is a short (1-2 pages tops) document describing what the team has accomplished and the immediate plans. I would like to see a brief outline of a proposed lecture in there, but beyond it, this is a progress report.

Followup. I will meet with each team during the week of February 14 to discuss the progress and the lecture plans. The meetings will take place in the same way as reading list approval meetings: during office hours, or by a separately arranged appointment.

Lecture Preparation.

Due Date: February 1 for lecture notes. March 8 and March 10 for lectures.

Each team will have about 35 minutes to actually deliver a lecture to the class. However, 35 minutes is too short for an actual introduction to any topic. Because of this, your lecture preparation will be more extensive.

¹E.g., it is ok to list a wikipedia page on a specific algorithm as reading material, but this needs to be backed up by the actual paper/textbook chapter discussing the algorithm in question.

Lecture Notes. Each team shall prepare lecture notes for a 2-3 hour introduction into the topic it has been assigned. That is, if someone were to teach out of these lecture notes and cover all the material contained in them, it should take them no less than 2 hours (and ideally - up to 3 hours, although if you want to prepare more extensive notes - more power to you!)

The lecture notes shall be prepared in a style similar to the instructor's lecture notes. They should be formal and contain all necessary math and pseudocode for the topics covered in the notes. The lecture notes should be well-organized and present a *cohesive narrative* related to your topic of study (see comments below).

I strongly recommend the use of LaTeX to prepare the notes. The simplest approach is to create an OverLeaf project, and maintain it throughout your work on this assignment. Any static deliverables should be saved to PDF and submitted to me (it is possible that your LaTeX documents change between the submission time and the time when I get to reviewing), however I will also ask you to pass the OverLeaf project links to me.

Lecture Notes approval. I will meet with every team during the week of February 28 (Feb 28 – Mar 3) to discuss the preliminary version of the lecture notes. Scheduling will be done in the same way as above.

Lecture. While your lecture notes assume 2–3 hours of lecture time, your actual in-class lecture will be 30-35 minutes in length. Each team will decide how it wants to approach this, what material to present, and what to keep out.

The in-class lectures should have depth. They cannot simply be a PowerPoint-supported overview of applications. In your allotted lecture time you must provide **solid technical insight**.

You may support your in-class lecture by additional materials - presentation slides, Jupyter notebook demos, and so on. Your full set of lecture notes (preliminary versions!) will be made available to all students on Monday, March 7.

Final Deliverables.

Due Date: March 15.

Following the in-class lecture, each team will receive some final feedback from me on the content and the presentation. You will have until March 15 (Tuesday of the finals week) to revise your materials (lecture notes/any supporting materials) and submit your final versions of the lecture notes, and of any other supplemental materials to the instructor.

Your final lecture notes shall contain the bibliography of all your reading materials. Any reading materials available to you in electronic form should be placed into a single google drive directory and shared with the instructor.

GOOD LUCK!