

2023-24 Big Data Computing Class: Exam Guidelines

Scope of the Document

This document contains the guidelines for the “2023-24 Big Data Computing” class exam. Please read through the whole document **carefully**.

Preliminaries

Starting this year, the exam will consist of a seminar on a research paper chosen from a curated list of distinguished conferences and journals that align with the topics covered in the course.

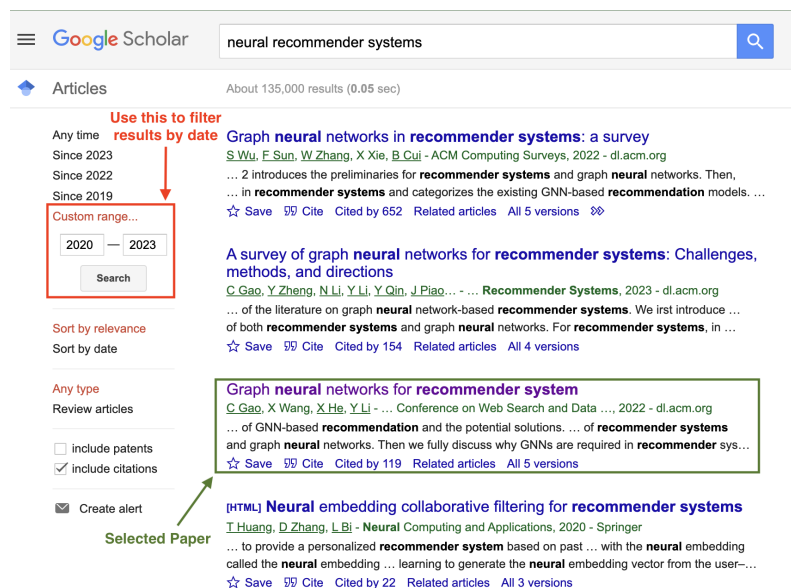
Since selecting a paper that simultaneously covers both units can be challenging, you can choose a research work that prevalently concerns one of the two units. For example, you can select work on big data architectures (first unit) or high-dimensional data representation learning (second unit).

Each seminar can be done **individually** or in a group of **at most two students**.

Paper Selection (Example)

Suppose you are interested in neural recommender systems. We suggest you go to [Google Scholar](#) and type in a search query using relevant keywords to your information needs (e.g., “neural recommender systems”).

As a result, you should expect to see something as follows:



Ensure that the paper you choose has a publication date **no earlier than 2020** (you can utilize the "Custom range" fields on the left column to filter results by date). Furthermore, the selected paper must be published in one of the two specified lists of venues, with each list corresponding to a specific unit.

List of conferences/journals covering topics seen in the first unit:

- USENIX Symposium on Networked Systems Design and Implementation (NSDI)
- ACM SIGCOMM
- ACM/IEEE International Conference for High Performance Computing, Networking, Storage, and Analysis (SC)
- EuroSys
- Conference on Machine Learning and Systems (MLSys)
- USENIX Annual Technical Conference (ATC)
- IEEE International Parallel & Distributed Processing Symposium (IPDPS)
- ACM International Symposium on High-Performance Parallel and Distributed Computing (HPDC)
- International Conference on Supercomputing (ICS)
- IEEE Transactions on Parallel and Distributed Systems
- Future Generation Computer Systems
- Journal of Parallel and Distributed Computing

List of conferences/journals covering topics seen in the second unit:

- ACM Conference on Information and Knowledge Management (CIKM)
- ACM Conference on Knowledge Discovery and Data Mining (KDD)
- ACM Conference on Recommender Systems (RecSys)
- ACM Special Interest Group on Information Retrieval (SIGIR)
- ACM TheWebConf (formerly known as WWW)
- ACM Conference on Web Search and Data Mining (WSDM)
- ACM Transactions on Information Systems (TOIS)
- ACM Transactions on Intelligent Systems and Technology (TIST)
- ACM Transactions on Knowledge Data and Discovery (TKDD)
- ACM Transactions on the Web (TWEB)
- IEEE Conference on Big Data (BigData)
- IEEE International Conference on Data Mining (ICDM)
- IEEE Transactions on Artificial Intelligence (TAI)
- IEEE Transactions on Big Data (TBD)
- IEEE Transactions on Knowledge and Data Engineering (TKDE)
- IEEE Transactions on Neural Networks and Learning Systems (TNNLS)
- IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI)
- Association for the Advancement of Artificial Intelligence (AAAI)
- International Conference on Artificial Intelligence and Statistics (AISTATS)
- International Conference on Machine Learning (ICML)
- International Joint Conferences on Artificial Intelligence (IJCAI)
- Conference on Neural Information Processing Systems (NeurIPS)
- Journal of Machine Learning Research (JMLR)
- Transactions on Machine Learning Research (TMLR)

Alternatively, you can directly explore the proceedings or articles on the websites of the abovementioned conferences/journals.

Using your Sapienza's institutional credentials, you will be able to access most of the digital libraries to download the paper you choose. More information is available at the following [link](#) (only in Italian, sorry!)

Paper Confirmation via Google Form

Please be aware that the lists of conferences/journals provided above may not be exhaustive. Therefore, if you opt for a seminar topic from a different source, ensure that your chosen paper is published in a top-tier venue. In either case, to communicate your choice, you must fill out an appropriate [Google Form](#) **at least one week before** the exam session date you plan to attend.

NOTE: *You are the only one responsible for the paper selection, and we will **not** provide preliminary approval/disapproval of the paper selection before the exam.*

Seminar Presentation Rules

Before illustrating the presentation rules, please check out the following guidelines on how to read and review a scientific paper: <https://2022-cs244.github.io/papers/L1-how-to-read.pdf>

The seminar is structured as follows:

- **15 minutes** for presenting the paper (if there are two speakers, divide the time equally between them)
- **5 minutes** Q&A on the presentation
- **5-10 minutes** Q&A on topics not covered in the presentation

Seminar Evaluation

If the seminar qualifies for grading, the final score will be determined based on the following criteria:

- Adherence to the Class Subject
- Complexity of the Topic
- Depth of Understanding
- Quality of Presentation

Exam Registration

We expect to organize **at least two** separate dates for each exam session. Before each exam session, you **must register on Infostud** for the appropriate date you choose. For example, for the upcoming January and February 2024 sessions, you can pick one of the following dates:

- 01/22/2024
- 01/31/2024
- 02/07/2024
- 02/14/2024

Dates for the June, July, and September 2024 sessions will be communicated later.

Best of luck!