How to do Research Project / Research Participation

Dongwoo Kim

2022/02/28

ML Lab

How to perform a research participation (CSED399)

- There are two different types of research participation
- If you are not familiar with machine learning (ML) and deep learning (DL), we can form a study group focused on DL
  - Study materials: List of Youtube videos
  - Coding materials: check the assignments in EECS498, Univ. of Michigan
- If you are familiar with ML & DL, choose a research topic presented in our lab homepage
  - MLLab homepage: ml.postech.ac.kr
  - Given a topic, I'll provide a list of curated research papers from easy to difficult ones.
  - Your job is to read & present these papers on a (bi-)weekly basis (you may present a single paper over multiple weeks based on your progress)

How to perform a research project (CSED499)

- I assume that students participating research project have enough background on ML & DL
- Note that the goal of CSED499 is to provide a chance to participate in real research to the students
  - Unfortunately, performing graduate-level research is not easy for most undergraduate students
  - Instead, we will follow a standard approach to conduct research with predefined milestones
  - At the end of each milestone, you need to present your results/ideas/findings to me
    - Send me an email to set up a meeting schedule
  - You don't need to follow the exact timeline but try to follow it

Research project pipeline

Choose a research topic

Read a couple of papers (not necessary to be the recent ones)

Find weakness or limitation of the papers

Look for ideas to improve the weakness and limitations

Implement your ideas

Perform required experiments

Summarize your results

Choose a research topic (week 2)

- Discuss with me to find a proper research topic
  - Based on your research participation history and interest

- From the topic, we can together curate a short reading list
  - 2 to 3 academic papers
  - From top-tiered conferences (NeurIPS, ICML, ICLR)
  - Not necessary to be the very recent papers

- Write a research proposal
  - Try focus on the background of the topic
  - Provide a rough timeline for the research project

Read a couple of papers &
Find weakness or limitation of the papers (week 3-6)

- Read the list of curated papers

- Critical thinking is important in this stage!
  - Do not believe what is written in the paper

- Find their limitations or weaknesses!
  - You don't need to find very complex limitations or weaknesses
  - Actually, it would be better if you can find a small idea that can address a minor part of the paper
  - For example, if you read about a paper using GAN (generative adversarial learning), many limitations of GAN, such as mode collapsing, is already identified by the others
    - Try to find as many limitations as possible and make a check list
  - Check that the paper you read fixes the limitations or not

Look for ideas to improve the limitation (weeks 6-8)

- Find approaches that addresses the limitation you have made in the previous stage
  - Starting from the limitation that you think the most critical
  - This may require you to read more papers!

- Summarize everything you read and found and be prepare to write an intermediate report from that
  - Summarize methods and limitations
  - Provide future plan (use this presentation as a guideline)

Implement your ideas (week 9-12)

- Implement the method that can potentially solve the limitations you provided
  - I recommend you to use github to maintain the source code
  - Make it available to everyone. This may give you a good reputation at some point in future

- Note that there are many existing implementations in the wild
  - Digging up Github or researchers homepage to find the original source code of the paper

Perform required experiments (1) (week 12-14)

- It's time to run the code.
  - I recommend you to use wandb to keep your results.
  - Check their excellent documentation (it is even available in Korean)
- If you need computation resources, please use CSED clusters
  - Check the CSED homepage to see the instruction

Perform required experiments (2) (week 12-14)

- Follow the training/validation/testing framework for all experiments
  - Model is trained on training set
  - Hyperparameter is tuned on validation set
  - Final performance is measured on test set
- Compare the performance of the original model and your model
- Don't panic if your result is not good enough
  - Bad results are also meaningful if you analyze the result well
- Try to analyze the performance
  - Plot training and validation loss curve: is it a problem of optimization?
  - If training loss is small and validation loss is large, check the regularization

Summarize your results (week 14-15)

- Organize the final report with the following five sections:

1. Introduction
   - Provide general background on the topic you cover. Minimize technical details.
2. Background
   - Describe the baseline algorithm of the paper you have read and discuss their limitations
3. Method
   - Provide the detailed method to tackle the limitations
4. Experiments
   - Show your results. Try to explain the results with evidence.
5. Conclusion