CS231A CA Session

Some tips for the course project

Chen Wang
2022/02/25

Potential project ideas

- Implement an existing method
- Benchmark for existing methods
- Propose a new method for a problem
- Application

1. implement an existing method

- Open-sourced method or Closed-sourced method
  - Both great! We will take the difficulty of implementing closed-sourced method into consideration for the grading.
  - Have a good estimation of time.
- Datasets
  - Start from the datasets mentioned in the paper and is easy to download and use. (Some datasets are huge. Some datasets do not exist any more.)
  - It is also good to start with your own dataset. But we should first estimate how hard is it to get the ground truth labels. We don’t recommend you to label the dataset by yourselves!
- What’s next step if I finished the implementation?
  - If it works very well. Congrats! More ablation studies is a good direction to improve the final report.
  - If it doesn’t work as the paper reported. Think about what might be the issue. Add some discussions/experiments in the final report.

2. Benchmark for existing methods

- How many methods do I need to consider?
  - It depends on the task. Recommendation is at least 3.
  - Consider open-sourced methods to compare with. If we want to implement some methods, we should have a good estimation of the time we need.
  - Focus on the analyzing the difference between the methods.
- How to find a good benchmark dataset?
  - Start with the ones mentioned in the paper. For benchmark projects, maybe we only need the evaluation set instead of the entire training set. (Some dataset has huge training set.)
  - It’s also good to build your own benchmark dataset but we need to make sure the accuracy of the ground truth labels.
- What should be the key part of a benchmark project?
  - Discussion and ablation study should be the key in the final report. Key question: What really matters for this problem?

3. Propose a new method for a problem

- Where should I start?
    - Based on one/two prior works is a good way to start! Key: Try to find the part that has potential to improve.
- How much should I improve?
    - Any improvements would be great! We are not looking for huge improvements. Add more discussion in the report: Why your propose idea improve the result?
- How many baselines should I consider?
    - >=1 should be good. Depends on the time we have.

4. Applications

- I’m running late. There are too many things to do for real-world applications!
  - We totally agree! This is why we recommend to have a good estimation of time.
  - Focus on some key components.
  - Keep in mind: it is hard to deploy a learning-based method on phone or some portable devices. Pay attention to the inference speed. Sometimes network compression is good.
  - Discuss about the challenges in the report. How we solve it?

Questions