Instructions for PhD Students

A presentation to my PhD students

Goals of PhD Students in Databases

- THE **CAKE**
  - ICML, NeurIPS, ICLR, KDD
- THE *ICING*
  - AAAI, IJCAI, WebConf (because sometimes the cake is not enough)
- Submission is most important
  - I will determine if the paper is good enough for submission.
  - Acceptance is often random – although it increases significantly your chances to find a job, it does not necessarily make you a better person (or scientist).
  - Rejection is part of the game – do not start questioning the meaning of life. If you are lucky you will get rejected again in the future.
- Acceptance builds CVs
- Rejection builds men

How to publish at ICML

- Have the following thought exercise:
  - Ask yourself how other people can benefit from your paper?
    - They can apply your theory to new tasks
    - They can apply your algorithm to their data
    - They can derive tricks you used in other domains
    - They can derive insights from your theorem
  - Examples of good benefits from your paper
    - You proposed a theoretically-sound loss function, init method, architecture
      - People directly inherit these tricks for their tasks
    - You proposed relevant problem and proposed a good algorithm for it
      - People can this problem with other data sets, invent commercial applications, or propose better algorithms to this problem
  - Example of bad benefits from your paper
    - You described that some algorithms or datasets are bad
      - Except for this knowledge, people don't gain anything in exchange → offer them better algorithms/datasets
    - You tested gazillion of models and found they are outperformed by baselines
      - Quite hard to sell; people like positive results; quite useful in practice but considered to be not science

Learn rules of the game

- When you become successful research scientist you will know rules of the game

- The quality of results that you obtained is not everything.
  - You need to sell your paper
    - Make the paper readable
    - No grammatic mistakes → proofread (e.g. Grammarly)
    - Make great visualization (takes a lot of time ~ weeks)
  - Luck is important: nowadays having 1 weak review with 2 strong accepts result in a reject.

- When you get rejection, don't question meaning of life
  - Decide on the next venue
  - Incorporate feedback (it's a gift to you)
  - Resubmit until it gets accepted (some papers take years to get published).

- After it got accepted, you need to promote

Learn rules of the game

- When you become successful research scientist you will know rules of the game

- After it got accepted, you need to promote
  - Write an intuitive blog post
  - Promote via social networks (reddit, twitter, medium)
  - If results are cool, try to find newspaper to highlight it

How to publish at ICML

- Have the following thought exercise:
  - Ask yourself how other people can benefit from your paper?
    - They can apply your theory to new tasks
    - They can apply your algorithm to their data
    - They can derive tricks you used in other domains
    - They can derive insights from your theorem
  - Examples of good benefits from your paper
    - You proposed a theoretically-sound loss function, init method, architecture
      - People directly inherit these tricks for their tasks
    - You proposed relevant problem and proposed a good algorithm for it
      - People can this problem with other data sets, invent commercial applications, or propose better algorithms to this problem
  - Example of bad benefits from your paper
    - You described that some algorithms or datasets are bad
      - Except for this knowledge, people don't gain anything in exchange → offer them better algorithms/datasets
    - You tested gazillion of models and found they are outperformed by baselines
      - Quite hard to sell; people like positive results; quite useful in practice but considered to be not science

Motivation

- You are mainly responsible for it
  - I already have >100 papers. +1 does not make a big difference in my CV – it makes a huge difference in yours.
  - I am a full professor – nothing makes a difference in my CV anymore.
- You have to come after me – not the other way around
  - I am tired of trying to motivate students.
  - You have to compete with the other students for my attention. I can always find one to work with anyway

How to Find and Keep Motivation

- Motivation is the most important qualification for a PhD student
- Read the recent ICML, NeurIPS, ICLR Proceedings
- You can ask to present material that you find interesting
- Initiate discussions
- Get involved in other (especially older) students' work
- Always try to write something – some of the best ideas will come to you by writing
- Writing means ***clarifying your mind***
  - a draft is something concrete – otherwise you may have done nothing as far as I am concerned.

How to Present Ideas (1) – for oral presentations and papers

- Only present topics that you understand CLEARLY
  - The most important problem is that sometimes people "do not understand that they do not understand".
  - If you do not understand a point, do not be afraid to admit it.
  - Understanding that you do not understand, is the most important step towards understanding.
  - Once you understand something it becomes simple
    - in which case you should be able to present it clearly
    - if it still seems complicated, probably you still do not understand it well enough

How to Present Ideas (2)

1. Describe the problem in detail
   - Most students fail at this point and the rest of their presentation is useless
2. Present the related work
3. Give the abstract idea of your solution
4. Explain why it is better than previous work
5. Only if steps 1-4 are clear go into the details
   - Most people are not as knowledgeable as you think they are.
   - Do not miss the point among the specifics of the solution.
   - A presentation is like a class – if the audience does not understand it is your fault, not theirs.
   - Present only things that you understand clearly – if something still looks complex, skip it.

How to Present Ideas (3)

- Be very careful about the notation
  - a good notation can make your life easy.
  - a bad notation is inexcusable. Given that most of you make a lot of English errors (excusable), you should make sure that at least your notation is carefully chosen.
- Spend a lot of time on good examples
- Read a lot of our previous papers and try to learn the writing style
  - it has been very successful

How to write papers

- All the above about presentations, plus
- LOVE your papers
  - Papers are forever.
  - LOVE shows in a paper.
  - Even if a paper gets rejected, eventually it always makes it somewhere.
  - Presentation is as important (if not more) than the actual work. Presentation is easier to improve.
  - Read and refine your draft as many times as possible. Then repeat the same process again (and again).

PhD Thesis Topic

- The topic is not important for the quality of a thesis
  - a really good student will produce results in any topic.
  - but a hot topic may help you find a good job later.
  - a topic in the supervisor's area is likely to increase your chances (in case you get stuck on your own).
- You do not necessarily need a concrete topic for your PhD thesis
  - Some of the best theses I have seen are collections of papers on a general topic.
  - If you are good enough assembling a thesis should take a few days of copy/paste work.
- Write the papers first and decide the topic later
  - The most difficult paper is the first one – the remaining ones usually come easy.
  - Normally, **the problem and not the solution** passes the paper. Try to think of innovative problems.

About Laziness

- It is very easy for PhD students to get lazy
  - they do not have office hours.
  - the supervisors are sometimes lazy.
- It is very easy for Professors to get lazy
  - they do not have office hours.
  - often they have tenure.
- The less you work, the less you want to work
  - which creates a vicious circle

About “Stupidity”

- Have not found yet an accurate definition:
  - Lack of common sense
  - Repeating the same mistake many times
- The really stupid person does not know that he is stupid
  - if he knew he would not be
- Intelligence (and competence) has limits. Stupidity has none
  - When I think that I have seen everything, there is always something new to surprise me.
- I will be the judge of your stupidity.

About Democratic Procedures

- The supervisor is always right (at least as far as you are concerned).
- We follow democratic procedures
  - You are encouraged to give your opinion.
  - I will decide.
- Your opinion is not very important before your first ICML paper.

Uncountable Qualifications of a PhD Student

- Motivation
- Programming Skills
- Background (e.g., papers read, mathematical background).
- Writing and Presentation Skills
- Creativity (ability to come up with new and interesting problems)
- Durability (“does not crack under pressure”)

Countable Qualifications of a PhD Student

- You are as good as
  - the number of **CAKE** papers (ICML, NeurIPS, ICLR, KDD) that you have co-authored.
  - the number of *ICING* papers *(*AAAI, IJCAI, WebConf) that you have co-authored.
  - most important are the papers where you are the first author.
- You get bonus points if I can count on you for
  - good reviews.
  - class tutorials, conference presentations.

Concrete Advice

- Read papers
- Bother me with questions, ideas etc.
- Ask to give presentations
- Write drafts of your ideas, or summaries of the papers that you have read and seem important
- Never miss a deadline
  - e.g., if you have a task that requires 1 week, ask for 2 to be on the safe side.
- Pay attention to the detail

Conclusion

- The ICML deadline is the end of January.
- During your studies you will have on the average 3 ICML deadlines.
- Missing one is important.
- I will be here to help and guide you.
- Do not crack under pressure.