

CS197 Harvard: AI Research Experiences

Fall 2022: Lecture 19 – “The AI Ninja”

Making Progress and Impact in AI Research

Harvard CS 197 Instructed by Pranav Rajpurkar. Guest lecture by Jia-Bin Huang. Website <https://cs197.seas.harvard.edu/>

Notes compiled by Sun-Jung Yum.

Abstract

Today, we are excited to welcome guest lecturer Prof. Jia-Bin Huang who shares key skills that are necessary in gaining expertise with AI research. These skills particularly cover two questions: (1) How can I make sure that I am making steady progress on my research? (2) How can I maximize the impact I have with my research work? You will learn about skills acquired from a long career in AI research, and how you can apply them as an undergraduate interested in AI research. I (Pranav) then ask Jia-Bin some of his thoughts on some of the common questions early career researchers face.



DALL-E Generation: “Becoming an artificial intelligence ninja”

Learning outcomes:

- Learn how to make steady progress in research, including managing your relation with your advisor, and skills to develop as an early career researcher.
- Gain a deeper understanding of how to increase the impact of your work

Introduction

Delighted to have this guest lecture from Prof. [Jia-Bin Huang](#).

From here on, Prof. Jia-Bin Huang is referred to in the first person ("I," "my," etc.)

Looking at this course, it's like a roadmap of how to "train yourself on how to be an AI ninja." Along the way, you learn tools (for SWE, Python/PyTorch, Transformers), methodologies (ideas, experiments, evaluation), papers (abstract/intro, method/results, figures), and dissemination (presentation, deploy model). This lecture is trying to address any gaps that were found here.

Specifically, we have two things. The first relates to the progression from methodology to papers. You learn all these tools on how to come up with ideas, and this seems like it might be a straightforward pass but in practice, it is very interactive. But, how do we make steady progress? This is something we will address in today's lecture.

The second part is the progression from papers to dissemination. It's not just about the papers. Once you have a paper, you want to make an impact on the world and think about how others will adapt and build upon your work. The question is, how do we create more impact? We'll address how we can create more impact once you have the work done.

Making Steady Progress

Mentor/Advisor Relationship

The question is, how do you make steady progress? There are two main components. One is the relationship between you and someone more senior, like a mentor, advisor, or a more senior graduate student. You want to make the most out of this relationship. The second component is the research skills that you want to learn and develop through this process, which we will dive into in the next section.

Point 1: Your mentor/advisor is an input-output machine.

Let's talk first about your mentor or advisor. There are a few points here. Your advisor is an input-output machine. When I was a graduate student, I wanted to impress my advisor. So, I treated my advisor as an input-only machine, thinking that I should do everything and just report the final results. But, in the end, I didn't really benefit much from this interaction. On the other hand, a lot of students often treat their advisor as an output-only machine, only doing things that their advisor tells them to do. But, this doesn't allow you to explore or lean on your own.

What you need to do is really treat this relationship as an input-output relationship. You want to frequently update your advisor on what you've done, and then get frequent feedback, in order to then form a good positive feedback loop.

Point 2: Show your work

Secondly, you always want to show your work. Often, students will make the mistake of showing only successes, feeling embarrassed to show any failures. But, you want to show your work, not specifically your successes, or your failures. In particular, when you show your work, you want to ask yourself a few questions: (1) What did you do? (2) Why did you do it? (3) How did you do it? (4) What did you find? (5) Why do you think this makes sense (or, on the contrary, why does it not make sense)? You want to think about these questions before you present your work to a mentor or advisor.

Point 3: Present failures

We talked about now just presenting successes, but also presenting failures. A lot of students will just say, "I tried this, and it doesn't work." But, this mindset does not work. Your advisor wants to help, but your job is to help them help you. So, how do you present failures?

Your advisor doesn't just want to see that something worked out of nowhere. What they want to see is the process. The following quote from "How to do research" by Bill Freeman, a professor at MIT, explains this very well:

"I've narrowed down the problem to step B. Until step A, you can see that it works, because you put in X and you get Y out, as we expect. You can see how it fails here at B. I've ruled out W and Z as the cause." – Prof. Freeman

As an advisor, you want to see something like this. This is the gist of research. When something doesn't work, you have to be able to dive in and digest, and then gradually narrow down the problems. This is all important research progress. You want to present this to your mentor and advisors so they can best help you to succeed.

Point 4: Provide contexts

Oftentimes, your advisor may be very busy, especially if they are senior faculty. You want to treat your advisor as a goldfish. When you step out the door, they likely will forget everything you just said. You want to maintain a detailed meeting minute on what you have done, what you will discuss, what's the plan, all in written form. Ideally, you should send them an email so that everyone is on the same page. And, lastly, whenever you discuss anything with your advisor, you should start with "why." Once you know the "why," everything will be easier to follow.

Junior students will often come in and just present what they did, and what the result was. But, without the context, they can't help you as much.

Point 5: Set expectations

Research can be very stressful at times. You want to set a boundary so you can maintain your mental health and have a work-life balance. Instead of saying, "I will finish this as soon as possible," you should set a date and time target for a specific task. Establish the fact that you will be busy during certain periods (like exam periods, etc.). This will set up mutual trust that you are staying on top of your responsibilities.

Research Skills

So, now, you have a strong, helpful relationship with your mentor/advisor. Now, let's talk about how you can improve your research skills. What kind of research skills can you develop as a junior aspiring researcher?

Point 1: Imagine success

The first thing is to imagine success. As a junior student, you typically don't have a very clear idea of what a good research problem is, or what the barriers to said problem are.

Usually, you define your research project as a kind of road map. You have a starting point, and an ending point. You want to develop your research so that you know where you are and you will reach your goal.

The very first thing is that, rather than starting with that immediate next step, you should visualize the goal. What is your goal? Is the goal exciting or not? This is the awesomeness test. Forgetting about technical difficulty, imagine that everything goes well and you achieve your goal. Is that awesome or not? If it is, you've passed the test. If it is not, then you shouldn't do this project. This will save you so much time from diving into a project that is not important, or does not lead to something interesting that you can then learn and grow from.

There is an article that describes how a famous mathematician at Bell Labs, Richard Hamming, "trolled" his colleagues. He would sit at a table with a group of chemistry scientists, and would ask them, "What are the important problems in your field?" And, they would talk about this. The next day, he would come to the same lunch with the same researchers and ask, "What are you working on?" And the next day, he would say, "If what you are doing is not important, and if you don't think it is going to lead to something important, why are you at Bell Labs working on it?" He was basically a bully, and he wasn't welcome after that. No one wanted to eat lunch with him. But, he did have a point. "If you do not work on an important problem, it's unlikely you'll do important work."

So that's the end goal. You want to think about your end goal, first.

Point 2: Work backward

Now, you've defined your end goal. But, how do you get started? Let's say you have your project roadmap, with A as your starting point and E as your goal, with milestones B, C, and D along the way. Most students will take a forward approach, going from $A \rightarrow B \rightarrow C \rightarrow D \rightarrow E$. This is very natural. You have some sort of data at A, and you need to implement something to get to B, from which you will need to make changes to get to C, etc. This is the forward model.

But, what I've found over the years is that this is really hard to do because you only achieve what you want to see, and already planned to see at the end.

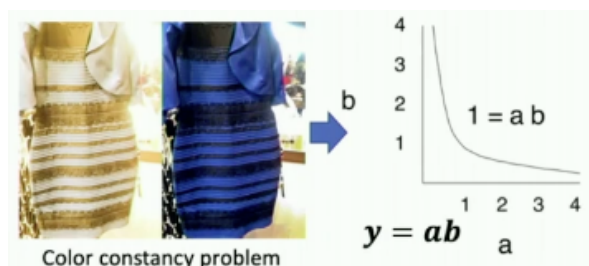
What I suggest is to follow this backward model. Let's break it down. You have some task between $A \rightarrow B$, $B \rightarrow C$, $C \rightarrow D$, and $D \rightarrow E$. Now, you work backward. You first assume that you have already solved A, B, and C. Imagine that you have a perfect input for the task $D \rightarrow E$, and work on that. Once you solve that, you go backward to $C \rightarrow D$. How can you get that perfect input for $D \rightarrow E$?

Why do we do this? First, you get to see the final outcome early. You see what you want to achieve much earlier, which keeps you motivated. You can see whether it's going to be awesome or not. Second, you can see the upper bound early. If you say you have the perfect input for $D \rightarrow E$, with no noise at all, you can get your upper bound on the best you could possibly do. If the upper bound is not awesome, you shouldn't do this. Third, because you are decomposing these tasks, every time you are working with something, you are working with perfect inputs. This allows you to focus on each task fully. Fourth, you can understand what is needed to succeed.

Point 3: Toy example

When you have a complex system, and a lot of the research you do is very complicated, it's useful to find The Simplest Toy Model That Captures The Main Idea (TSTMTCTMI). What is that? Let's use two examples.

First, let's think about the color constancy problem, where when you observe a pixel color, you want to see how you can factorize this color into something that's the illuminated color and the reflectance of the color. Like, what's the color of this dress? You can study this from a very simple, toy example, such as $y = ab$. This very abstract, simple model captures the simple idea. And, then, you can develop all the intuition here.



As a second example, let's consider 3D video. Given a video of complex motion, how do I create a 3D video? Instead of thinking about this, we come up with a toy example. For example, we consider a boy that is simply moving constantly in space. We can think through this toy model and the development there, and the rest will follow. This will make your research progress much, much more efficient.



Point 4: Simple things first

Let's say you want to build a computer vision-based sheep counter. A lot of the time, the first example they try is this:



Of course, it doesn't work. And when it doesn't work, it doesn't give you any insight. Instead, what you want to do is start with a simple example first, like this.



Then, you can really understand what is going on, and once this works, you can move on to more complicated examples. You want to keep it simple.

Point 5: One thing at a time

Not only should you keep things simple, but you should also try things one at a time. Consider the lecture on Hydra, where we worked with all this config. Let's say you try something like this, where you change your batch size, change your learning rate, add some loss or disable some loss...

```
batch_size = 4 -> batch_size = 16
learning_rate = 1e-3 -> learning_rate = 1e-5
use_abc_loss = False -> use_abc_loss = True
```

And then, oh no, your results are worse. What do you learn? Nothing. You don't learn anything by changing multiple things at the same time. What you want is to slow down, so that you can speed up. Change one thing at a time so that you know that the output is something that is the effect of the change in input.

Creating More Impact

Say that, through the process described above, you now have the tools, you have the methodology, and you can progress towards completing a research project. But, how do you create more impact from there?

Let's start with a personal story. I was presenting a virtual poster at ECCV 2020, a computer vision conference. Over the course of 2.5 hours, 5 people visited my poster. I was obligated to do that, so I stood there. Later, I put together a video of the results of my work and shared it on Twitter. It got 508,000 impressions. To reach that same level of visibility, I would have had to stand at my poster for 29 years! You can create much more impact if you put your work out there, and there are a couple of solutions for this.

Point 1: Pick a good name

For one, you want to pick a good name for your work. Consider this [talk](#) given by David Patterson from Berkeley and Google.

Patterson says that your acronym should be three or four letters, because it's very hard to remember a phrase if it is more than three or four words. And then, you throw a bunch of words up on the board that match the three or four letters that describe the things you are doing and are a word in English. And, vowels are pretty important.

The general recipe is that you want something that people can memorize.

Point 2: Make all your results available

If you put your paper behind some sort of paywall, for example, people aren't able to download it. But if you make it available and let people download it, people can share it. It's all about making your work more easily findable and accessible.

Point 3: Lower the barrier for others to follow

Once you have done some work, you want to make sure that you're actively convincing people to build upon your work. You can publish a simple Github page to release your model, you can use Google Colab so that there are fewer technical barriers, you can use Hugging Face for machine learning models, and you can use PyTorch Hub to allow for people to download your model very easily. All of these are strategies to disseminate your work better.

Point 4: Make others' life easier

When you did this project, you went through steps A, B, C, D, and E in your roadmap. Along the way, you may have a lot of things that are variables for other people. For example, you may implement some baseline. Or, you may have an evaluation script, or data preprocessing, or a whole new dataset. All of these things allow for greater impact for your work. You don't want to waste this. Don't just simply present the final product.

Point 5: Show your work!

One of the most important components is that you want to show your work. You can show your work on your website, or your YouTube, or Twitter. As a researcher, you, yourself, are your brand. You want to nurture your brand, and build your audience. This allows you to create more impact.

Even if you don't have any scientific papers published, it's not about showing that work you've done. It's not just about self-promotion, it's about self-discovery. It's about the discovery of the process. For example, maybe you learned something from class and you want to present it. You can summarize it, and show it to others.

The process itself is actually very enticing for other people to learn; people can't relate as much to someone who is already a master. You can help others. Even as a beginner of anything, you can still build up your brand and share your learning process.

Discussion Questions

In this section, I (Pranav) asked Jia-Bin his thoughts on some of the common questions early career researchers face.

How do skills transfer into industry?

Question: A lot of people in the audience are in their undergraduate years and are thinking about going into industry. So the question is, as someone who has spent time in industry and academia, what skills that one learns in research transfer in industry and what skills are less important?

Answer: In my personal experience, I am mostly in the research arms within the labs, so I may not know much about the pure industry, pure product-driven life. But, in general, lots of research skills are very transferable to industry. For example, how you make progress is very similar to how you would think about, if you were in a product team, how you iterate a product, how you debug, how you set the goals. These are important skills in both research and industry.

Some people say that, in industry, you don't need to write papers or do presentations, you just need to write code. This is true, but you also need to be a good communicator. In industry, you see a lot of write-ups or whitepapers to convince others to invest more into your project. That requires skills to convince other people, and, those are skills that come from the question of, "how do you logically describe your impact," and "how do you plan to do these things?"

Should I go to graduate school?

Question: Let's say I'm a person who is deciding whether or not to go to graduate school. How should I make that decision?

Answer: I think the first thing you want to do is to try research. Try it once or twice, because you don't know whether or not you like it. Some people don't like this type of uncertainty, and some people really thrive from it. So, if you try it, and you don't like it, you can always go into industry. And, maybe you can still even do research within the research arms of industry. But the problem is that, if you don't try it, there is no way to know.

If you try it, it's possible that you really like it, and you find some projects that really interest you. Then, you can think about how you can pursue those projects more as a graduate student.

That's why, when you do a graduate school application, you have an SLP. You need to state your purpose, you need to be very crystal clear about your experience, and how you want to pursue your goals in the next five years, for example.

Research in industry vs. research in academia?

Question: Let's say I'm not sure if I want to do research in industry or in academia. Should I go to graduate school, or should I take on industry research experience?

Answer: There are some split opinions here. Some people say that you should go into industry first, and then you can understand what you like and don't really like. You get that real world experience, and when you go back to academia, you become much more driven with a much more clear goal.

Another option is to try research first while you are at school. If you don't like it, you have many more opportunities in industry.

How should I select a research problem?

Question: How can I, as a prospective research student, select a good problem to work on for research?

Answer: This is a challenging question. The easy answer is that you want to work with someone who can judge better, who has better "taste" than you. For example, your advisor or some research mentors. But, even when working with those mentors, you want to imagine success, like mentioned before. You want to visualize what it will look like if you succeed in a given project. What kind of new capability will you generate? What kind of new results will you discover? If that is exciting, that's probably a good problem to tackle. If not, then you don't want to go down that path.

It's an opportunity cost. You want to free yourself and your time up by rejecting mediocre ideas, allowing you to work on something more exciting and more important.

How should I select a research problem?

Question: You talked about how it's important to work backwards. How do you know when to pivot or when to stop something entirely? Let's say things aren't really working out. You're not finding very interesting results. Do you pivot? Do you try something similar that could help? Or, do you drop it entirely? How do you know?

Answer: It's always a tradeoff. Sometimes, you might work on a project and get stuck. But, you've already spent a lot of time on it, so you want to get something out of it. It's hard to give a clear answer; it's definitely a case-by-case issue. If you want to pivot, you want to make sure that you have a clear path. You can lay out that research path towards a goal and visualize that goal, and ask yourself, "Is that goal exciting, or not?" That's a good way of doing early rejection.

Also, by working backwards, you can quickly realize if something is achievable or not. If you have the best possible input for the last stage, and even that doesn't work, you know that it is not feasible. But if you are working using the forward model, you probably won't realize that until the end of your 6 months of effort. You'll waste a lot of time. You want to balance your opportunity cost; you want to weigh how you invest your time.

How can I approach a professor or a lab group?

Question: Suppose that you're interested in a research project, and you notice that a professor or a lab group is working on something similar. Where should I be in terms of thinking through my project to be able to then approach that lab and say, "I want to work with you"? Let's say a

student came up to you and said, "I want to do research in this area." What would you want to see from them to then want to take them up on their research project? What would make you think that a student is a good fit?

Answer: Taking on students as a professor or a research scientist is a sort of investment, especially for an assistant professor. You're asking them for their time, and time is the most valuable thing for them at this stage. So, you want to make sure that you are providing value. You first want to show that you are driven and motivated. You want to show that you're not just looking to get research credit in order to then apply for an internship, for example. In addition, when people look at junior students, they typically don't want an implementation robot. They want to see your thoughts. You want to demonstrate what kind of commitment you can put in, what other research experiences you have, and how those experiences might be relevant to the lab. Throw ideas out there, even if they aren't realistic or great. The pure fact that you have ideas is what is important.

The worst scenario for an advisor is that they will spend a lot of time working with you but they don't get anything out of it. For example, if there was no good project outcome. This is what makes faculty hesitant to take on students. It may also be that the commitment is not a longer-term commitment. Let's say you only want to work with them for 1 semester. In that time, it's unlikely that you will get anything substantial done, or achieve anything that is publishable that will help them and their career.

You want to understand from their perspective.

Thank you

To Prof. [Jia-Bin Huang](#) for this guest lecture. Prof. Jia-Bin Huang has written thoughtfully [here](#) on making research progress, working with mentors effectively, presentation, communication, career advice, and productivity.