Machine Learning Interviews

Chip Huyen (@chipro)
chip@huyenchip.com

Slides posted on Twitter @chipro

My background

Writing

Product

Cốc Cốc browser
20M+ monthly active users

Baomoi.com
acquired by VNG

Youth Asia
acquired by Groupon

AI/ML

PRIMER

NETFLIX

NVIDIA

My work (very pro-OSS)

Doing cool stuff

Chip Huyen
chiphuyen

Edit profile

Deep Learning Algorithm **@NVIDIA**. Taught "TensorFlow for Deep Learning Research" **@Stanford**. Author of 3 bestselling books. Insta: huyenchip19

@NVIDIA

Overview | Repositories 52 | Projects 0 | Packages 0 | Stars 152 | Followers 4.1k | Following 34

Pinned

Customize your pins

lazynlp
Library to scrape and clean web pages to create massive datasets.
Python ★ 1.6k 234

stanford-tensorflow-tutorials
This repository contains code examples for the Stanford's course: TensorFlow for Deep Learning Research.
Python ★ 9.1k 4.2k

python-is-cool
Cool Python features for machine learning that I used to be too afraid to use
Jupyter Notebook ★ 2k 288

NVIDIA/**NeMo**
Neural Modules: a toolkit for conversational AI
Python ★ 829 82

NVIDIA/**OpenSeq2Seq**
Toolkit for efficient experimentation with Speech Recognition, Text2Speech and NLP
Python ★ 1.1k 254

sotawhat
Returns latest research results by crawling arxiv papers and summarizing abstracts. Helps you stay afloat with so many new papers everyday.
Python ★ 1.1k 140

Contents

- Machine learning jobs
- Getting a job in ML
- Understanding the interviewers’ mindset
- Interview process
- Recruiting pipeline

Machine learning jobs

- Research vs applied research
- Research scientist vs research engineer
- Data scientist vs machine learning engineer

Research	Applied research
Find the answers for **fundamental** questions and expand the body of theoretical knowledge.	Find solutions to **practical** problems
Ex: develop a new learning method for unsupervised transfer learning	Ex: develop techniques to make that new learning method work on a real world dataset
Focus on **long term** outcome	Focus on **immediate** commercial outcome

Caveats

- Cutting-edge research is spearheaded by big corporations
- Lacking theories to explain methods that work well empirically

Research scientist	Research engineer
Develop original ideas	Use engineering to actualize these ideas
Might require PhDs	Don’t require PhDs
Might act as an advisor to research engineers	Springboard to become research scientist

Caveats

- Depends on organizations/teams. In some teams, there are virtually no difference
- Scientists & engineers can be equal first authors (e.g. GPT-2, Transformer paper)

Data scientist	ML Engineers
Extract knowledge and insights from structured and unstructured data	ML models learn from data -> ML is part of data science
Use data to help company make decisions	Develop models to turn data into products
Is a scientist -> engineering isn’t a top priority	Is an engineer -> engineering is a top priority

Caveats

- MLEs at startups might spend most of their time wrangling data, understanding data, setting up infrastructure, and deploying models instead of training ML models.

Machine Learning Engineer (MLE)

Umbrella term to cover:

- research engineer
- devrel engineer
- data scientist
- deep learning engineer
- generic machine learning engineer

It does not cover:

- ML DevOps
- framework engineer

ML at big companies
#
ML at startups

Big companies	Startups
Can afford research	Can’t (estimated cost of AlphaStar is $26M)

Big companies	Startups
Can afford research	Can’t (estimated cost of AlphaStar is $26M)
Can afford specialists	Need generalists

Big companies	Startups
Can afford research	Can’t (estimated cost of AlphaStar is $26M)
Can afford specialists	Need generalists
Standardized process	Make up process as they go

Getting a job in ML

Six common paths

1. BS/MS in ML -> ML engineer
   Tech Ivies -> FAANG/startups

Six common paths

1. BS/MS in ML -> ML engineer
2. PhD in ML -> ML researcher
   Published at top-tier conferences -> FAANG/ML-first startups

Six common paths

1. BS/MS in ML -> ML engineer
2. PhD in ML -> ML researcher
3. Data scientist -> on-job training -> ML engineer/researcher
   Companies want to start using ML

Six common paths

1. BS/MS in ML -> ML engineer
2. PhD in ML -> ML researcher
3. Data scientist -> on-job training -> ML engineer/researcher
4. Software engineer -> courses -> ML engineer
   Software engineers want to transition into ML

Six common paths

1. BS/MS in ML -> ML engineer
2. PhD in ML -> ML researcher
3. Data scientist -> on-job training -> ML engineer/researcher
4. Software engineer -> courses -> ML engineer
5. Adjacent fields -> on-job training -> ML researcher
   Not enough talents from AI/ML PhDs
   Ex: stats, physics, math

Six common paths

1. BS/MS in ML -> ML engineer
2. PhD in ML -> ML researcher
3. Data scientist -> on-job training -> ML engineer/researcher
4. Software engineer -> courses -> ML engineer
5. Adjacent fields -> on-job training -> ML researcher
6. Unrelated fields -> residency/fellowship -> ML researcher
   Big companies: Google, Facebook, MSFT, NVIDIA, etc.
   Ex: health, art, architecture, agriculture

Avoid anyone that promises you ML expertise in days or weeks!!

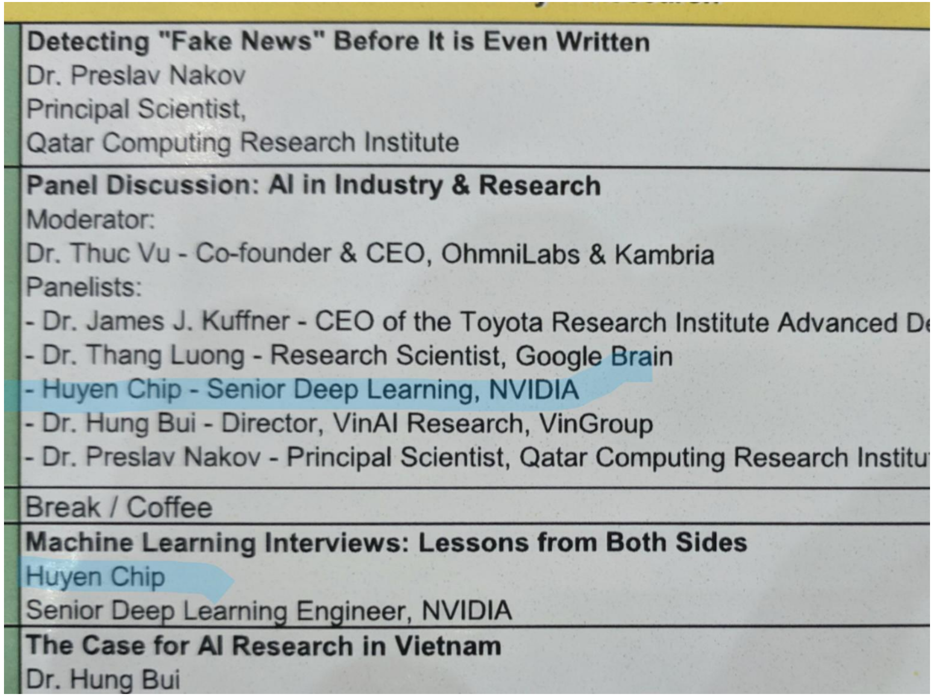
Senior role	Junior role
Hired for skills	Hired for attitude

Do you need a PhD?

[redacted], I feel like companies that I'm interested in such as Google and Lyft might ignore my application due to not having an MS/PhD

know fair amount of things in NLP and speech as well. Given that I don't have an MS/PhD, expertise in one field isn't taking me anywhere. [redacted] not required?

Do you need a PhD?



Detecting "Fake News" Before It is Even Written
Dr. Preslav Nakov
Principal Scientist,
Qatar Computing Research Institute

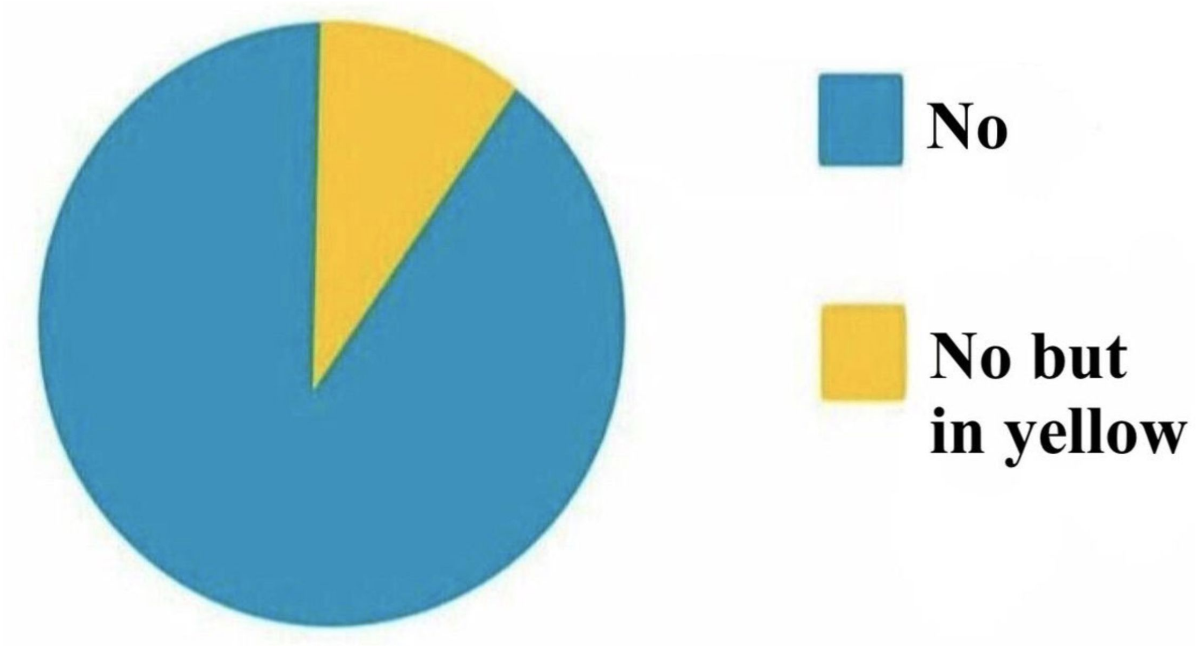
Panel Discussion: AI in Industry & Research
Moderator:
Dr. Thuc Vu - Co-founder & CEO, OhmniLabs & Kambria
Panelists:
- Dr. James J. Kuffner - CEO of the Toyota Research Institute Advanced De
- Dr. Thang Luong - Research Scientist, Google Brain
- Huyen Chip - Senior Deep Learning, NVIDIA
- Dr. Hung Bui - Director, VinAI Research, VinGroup
- Dr. Preslav Nakov - Principal Scientist, Qatar Computing Research Institu

Break / Coffee

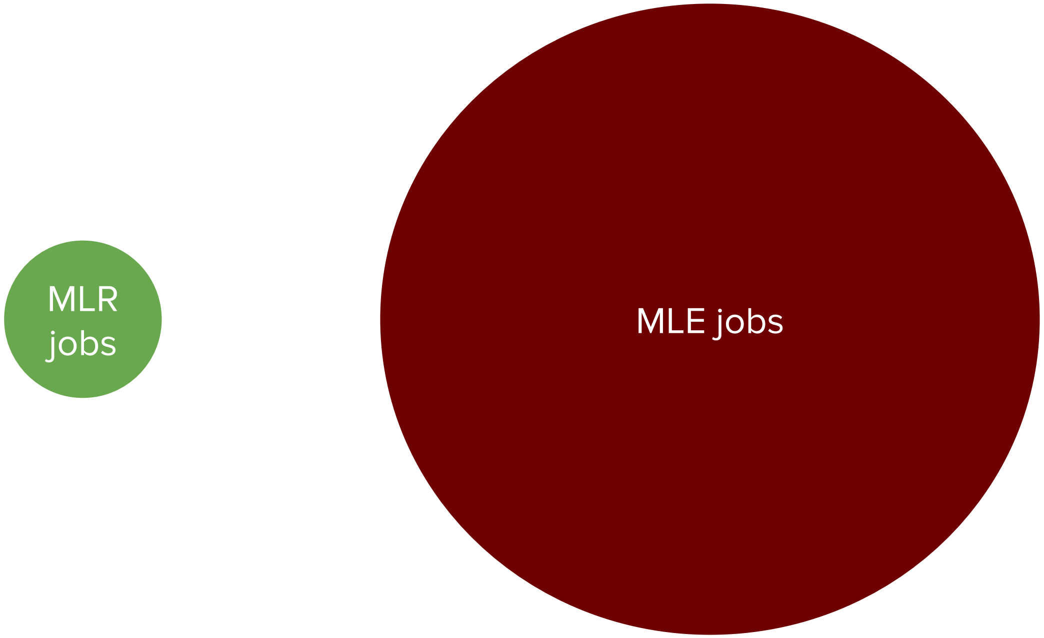
Machine Learning Interviews: Lessons from Both Sides
Huyen Chip
Senior Deep Learning Engineer, NVIDIA

The Case for AI Research in Vietnam
Dr. Hung Bui

Do you need a PhD?



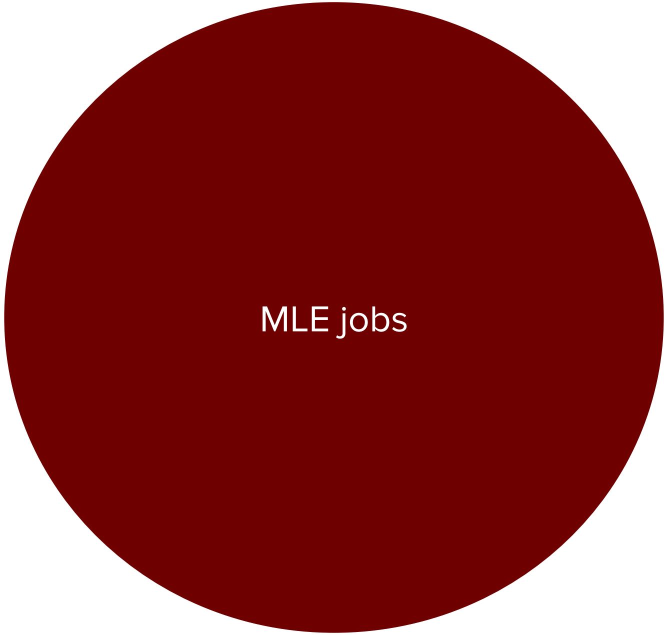
The only role that might require a PhD is (applied) research scientist



The only role that might require a PhD is (applied) research scientist

We need more engineers:

- **To improve research**
- **To productize research**



Understanding the interviewers’ mindset

1. Companies hate hiring

- Expensive for companies
- Stressful for hiring managers
- Boring for interviewers

2. Companies don’t want the best people

They want the best people who can do a reasonable job within the time and monetary constraints.

3. Companies don’t know what they’re hiring for

They don’t even know for sure if they’ll need that person
-> job descriptions for reference purposes only

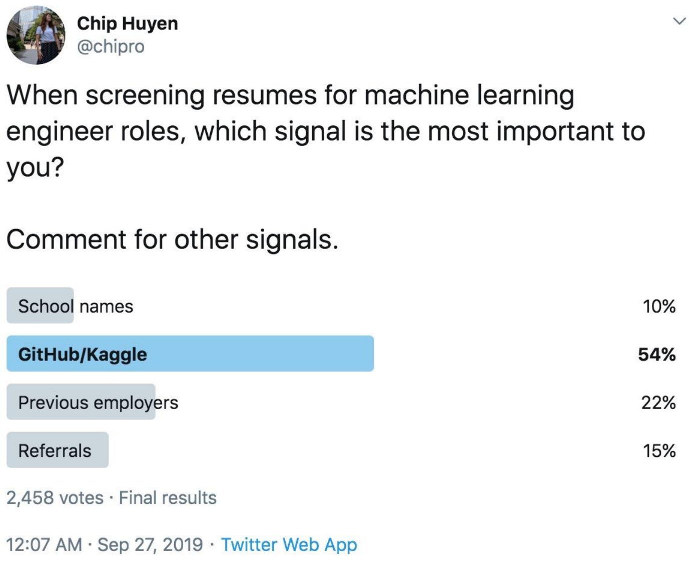
4. Most recruiters can’t evaluate technical skills

They rely on weak signals:

- previous employers
- degrees
- awards/papers
- GitHub/Kaggle
- referrals

Weak signals

- previous employers
- degrees
- awards/papers
- GitHub/Kaggle
- referrals



PSA: Past projects aren’t meritocratic

- Not everyone can afford to contribute to OSS or do Kaggle competitions.
- Placing too much importance on voluntary activities punishes candidates from less privileged background.

5. Most interviewers are bad

Little or no training for interviewers even at big companies

6. Interview outcome depends on many random variables

It is, in no way, a reflection of your ability or your self-worth

Interview process

Interview process

Evolved out of the traditional software engineering interview process

Interview process

1. Resume screen
2. Phone screen
3. Coding challenges / take-home assignments
4. Technical offsite interviews (1 - 2)
5. Onsite interviews (4 - 8)

Interview process

1. Resume screen
2. Phone screen
3. Coding challenges / take-home assignments
4. Technical offsite interviews (1 - 2)
5. Onsite interviews (4 - 8)

PSA:
Take-home assignments punish less privileged candidates!!

Onsites are tiring

10.00 - 11.00	Natalie Portman	Software Engineer	Coding question
11.00 - 11.45	Stephen Chow	Research Scientist	ML theory
11.45 - 13.00	Constance Wu	Engineering Manager	Meeting the team
13.00 - 14.30	Irrfan Khan	Research Engineer	ML implementation
14.30 - 15.00	Dave Chappelle	VP of Engineering	Behavioral questions

Bad interview questions

1. Questions that ask for the retention of knowledge that can be easily looked up
   "write down the equation for the Adam optimizer."

Bad interview questions

1. Questions that ask for the retention of knowledge that can be easily looked up
2. Questions that evaluate irrelevant skills
   “write a linked list.”

Bad interview questions

1. Questions that ask for the retention of knowledge that can be easily looked up
2. Questions that evaluate irrelevant skills
3. Questions whose solutions rely on one single insight
   “take derivative of x^x.”

Bad interview questions

1. Questions that ask for the retention of knowledge that can be easily looked up
2. Questions that evaluate irrelevant skills
3. Questions whose solutions rely on one single insight
4. Questions that try to evaluate multiple skills at once
   "explain PCA to your grandma."

Bad interview questions

1. Questions that ask for the retention of knowledge that can be easily looked up
2. Questions that evaluate irrelevant skills
3. Questions whose solutions rely on one single insight
4. Questions that try to evaluate multiple skills at once
5. Questions that use specific hard-to-remember names
   “Moore–Penrose inverse” or “Frobenius norm”

Bad interview questions

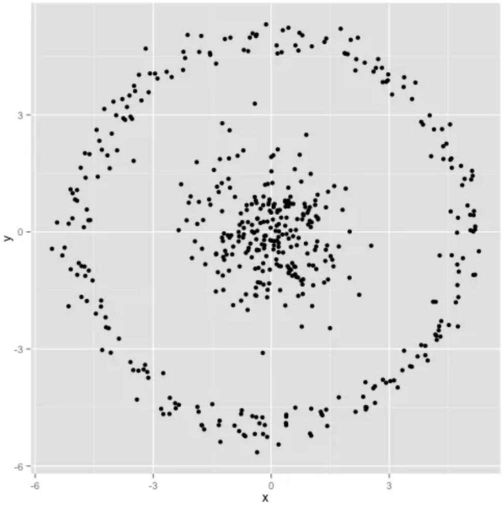
1. Questions that ask for the retention of knowledge that can be easily looked up
2. Questions that evaluate irrelevant skills
3. Questions whose solutions rely on one single insight
4. Questions that try to evaluate multiple skills at once
5. Questions that use specific hard-to-remember names
6. Open-ended questions with one expected answer

Bad interview questions

1. Questions that ask for the retention of knowledge that can be easily looked up
2. Questions that evaluate irrelevant skills
3. Questions whose solutions rely on one single insight
4. Questions that try to evaluate multiple skills at once
5. Questions that use specific hard-to-remember names
6. Open-ended questions with one expected answer
7. Easy questions during later interview rounds
   “find the longest common subsequence.”

Examples of good interview questions

Given the following dataset, can you predict how K-means clustering works on it? Explain.



Examples of good interview questions

Imagine you have to train a NER model on the text corpus A. Would you make A case-sensitive or not?

Examples of good interview questions

Duolingo is a platform for language learning. When a student is learning a new language, Duolingo wants to recommend increasingly difficult stories to read.

1. How would you measure the difficulty level of a story?
2. Given a story, how would you edit it to make it easier or more difficult?

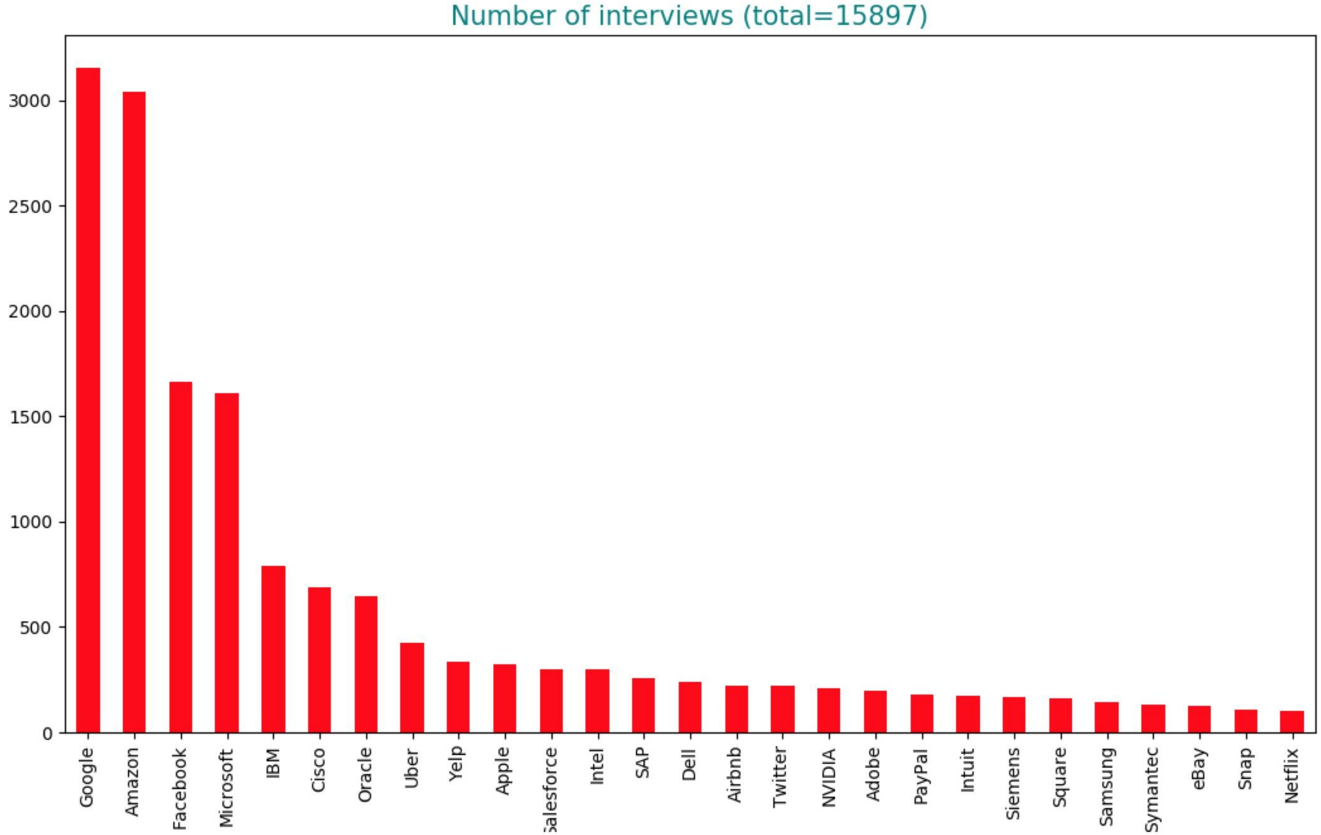
Alternative interview formats

- Multiple choice quiz
- Quiz
- Code debugging
- Pair programming
- Good cop, bad cop

Recruiting pipeline

1. Onsite-to-offer ratio
2. Offer yield rate

- **15,897** Glassdoor interview reviews
- **27** major tech companies



Software Engineer Interview

Anonymous Interview Candidate

No Offer

Positive Experience

Difficult Interview

- Result
- Experience
- Difficulty

Application

I applied through a recruiter. The process took 2+ months. I interviewed at Google in June 2014.

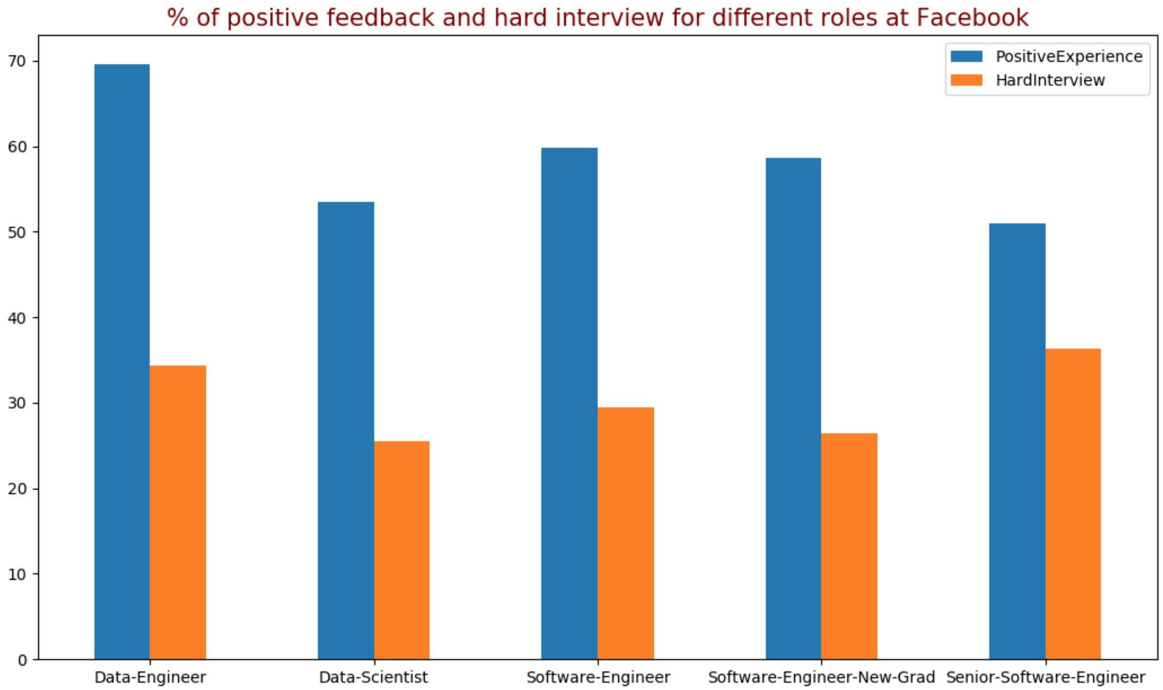
Interview

Was approached by a recruiter, we had an initial phone screen. From there, I had a technical phone interview with a SWE. The questions were pretty straight forward, nothing too difficult. After that, I had four on-site interviews. Two of them went very well, one went pretty well, and I did pretty poorly in the other. I signed an NDA, so I can't go into details, but my suggestion to...

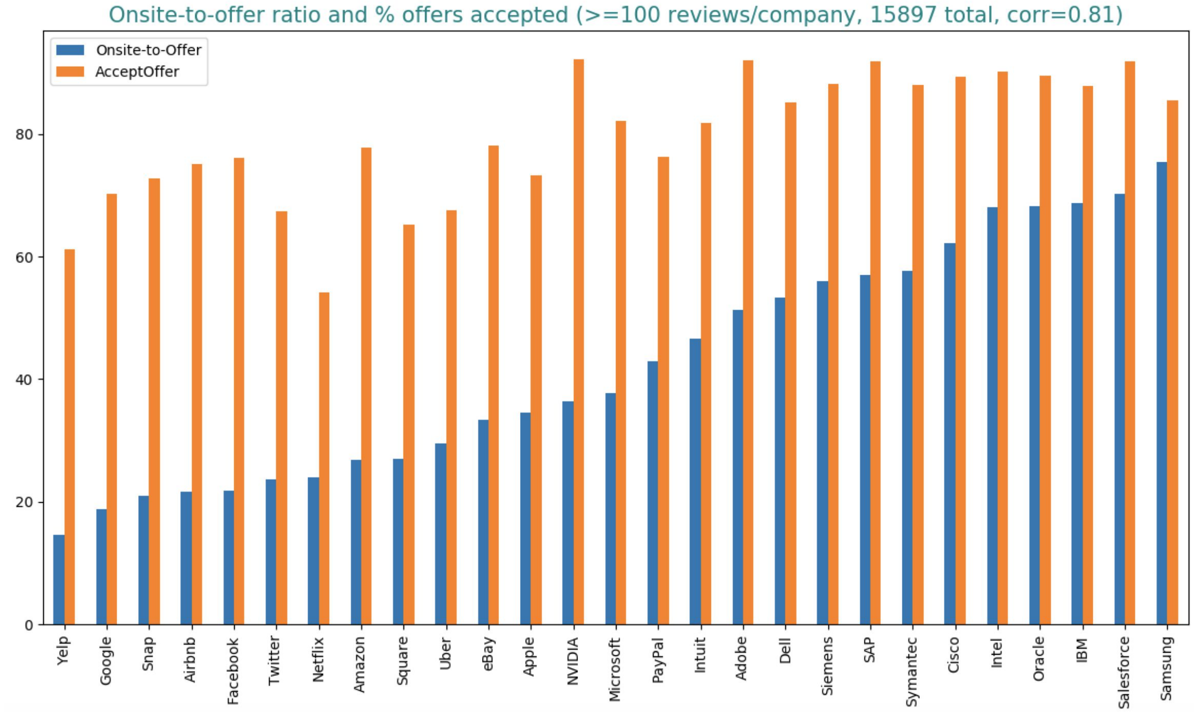
Biases

- few people leave reviews for anything online
- reviews are likely compelled by either a really good or really bad experience
- those who receive offers are more likely to give reviews than those who don’t
- those who accept offers are more likely to give reviews than those who decline
- …

Interview feedback varies wildly even within the same company

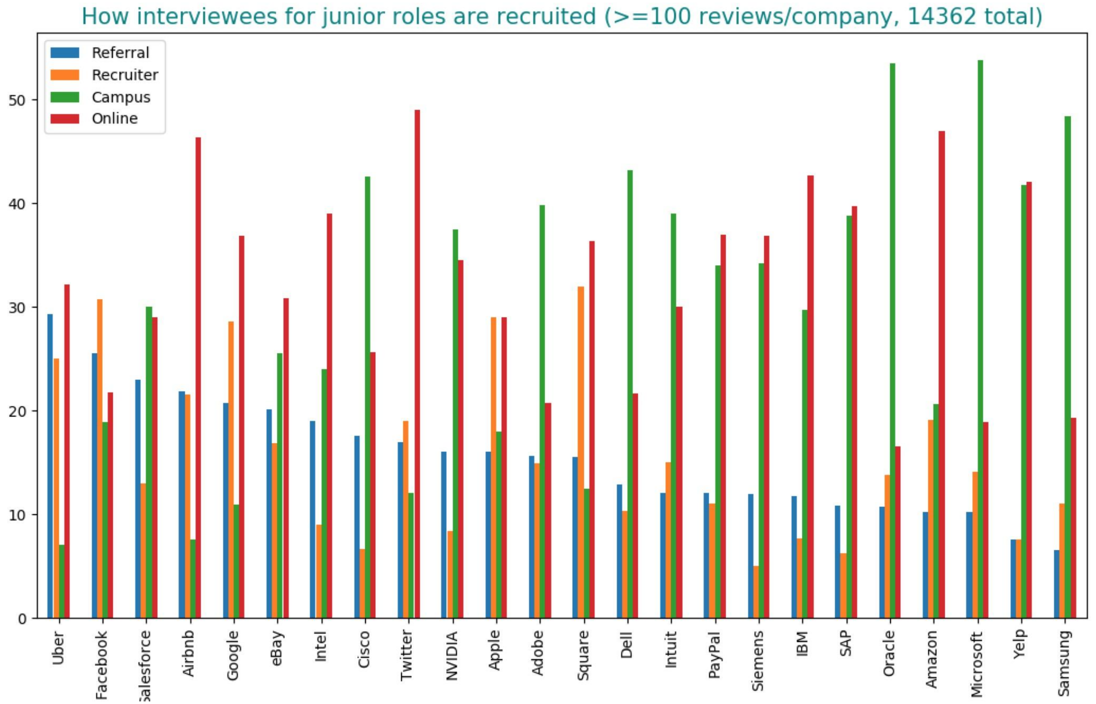


The higher the onsite-to-offer ratio, the more likely offers are accepted



Link to blog post

Most junior roles are sourced through campus or referrals



How important are referrals?

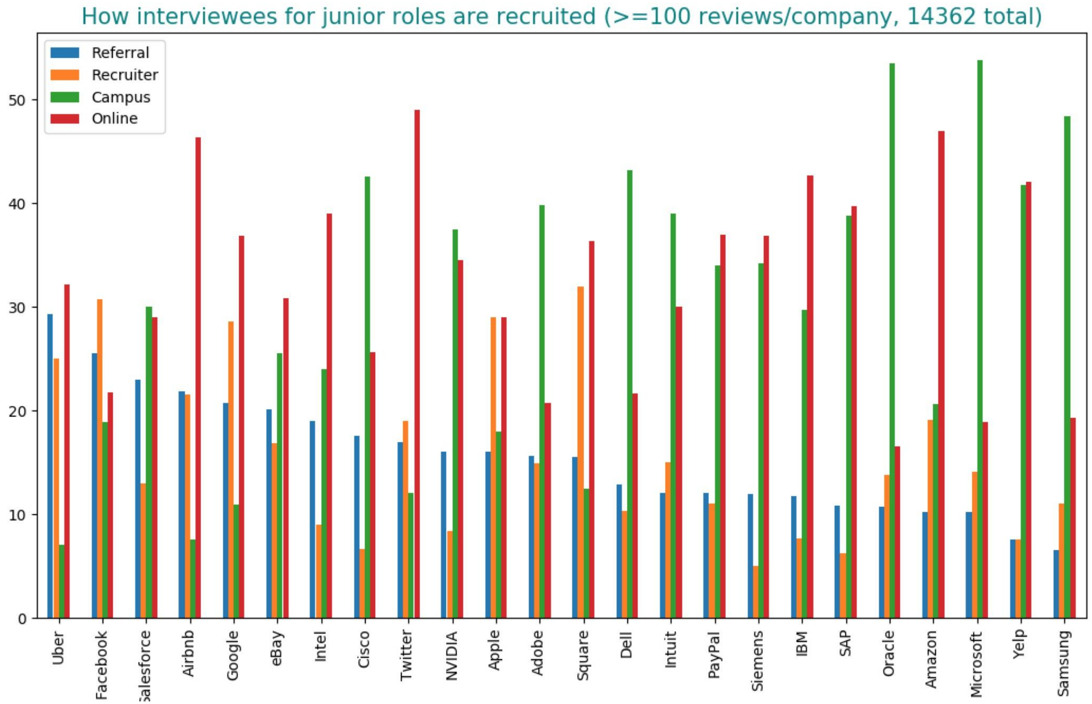
"By at least a 10x margin, the best candidate sources I've ever seen are friends and friends of friends."

- Sam Altman (Y Combinator)

"Referral quality was incredibly important – the 8 worst hires … were all unknown to me and everyone at the company at the time of hiring."

- Lukas Biewald (Figure 8, Weights & Biases)

Most junior roles are sourced through campus or referrals



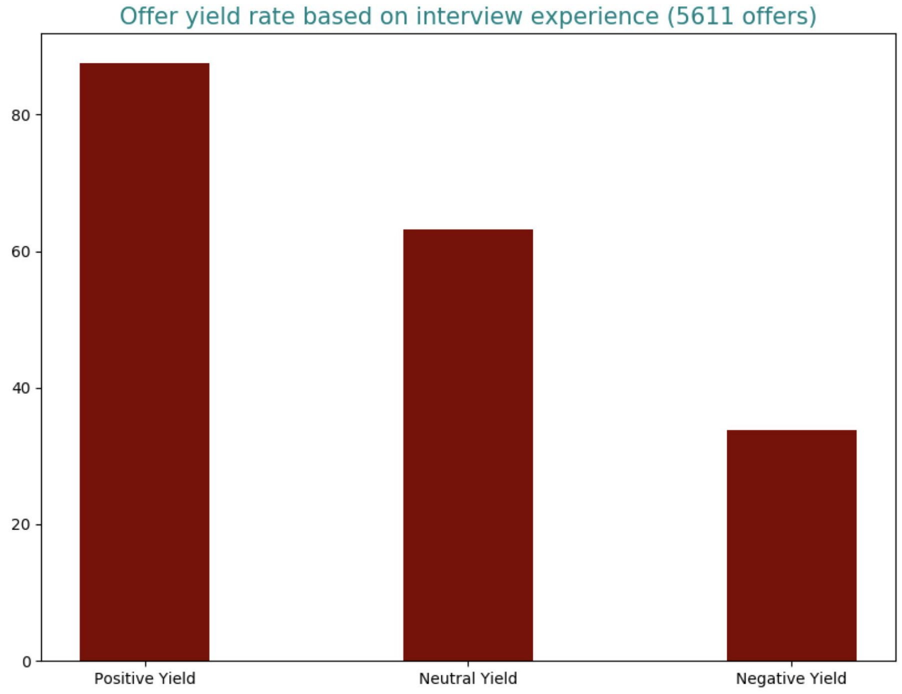
Tech Ivy mixer alert!!!

What if you don’t go to a fancy engineering school or know people who can refer you?

“Be so good they can’t ignore you”

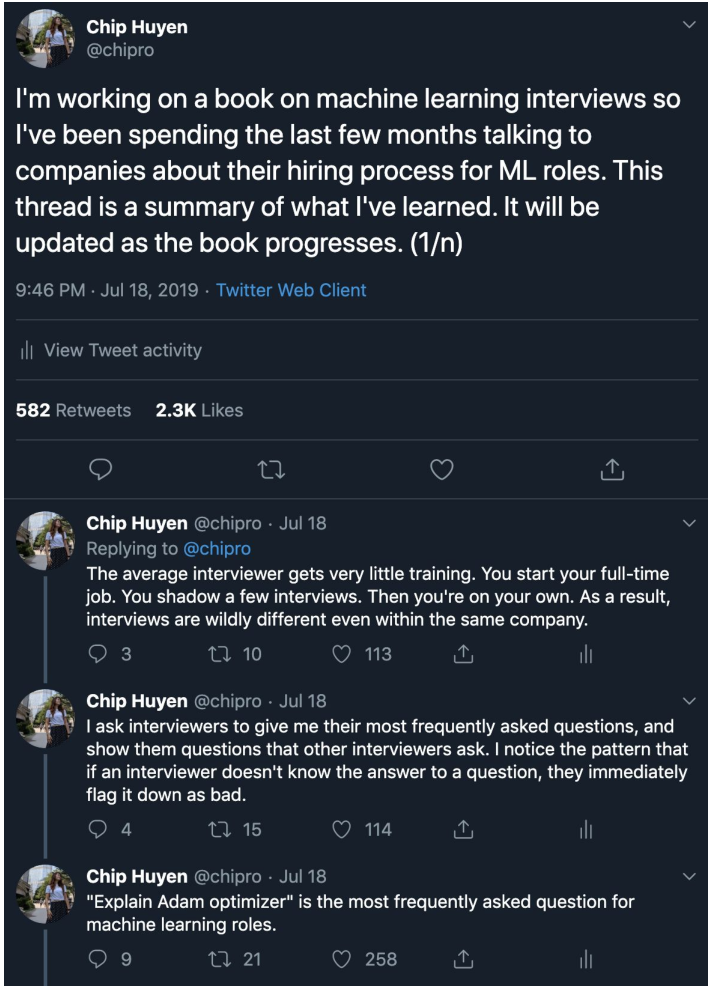
- Steve Martin

Candidates with negative experience are less likely to accept offers



General tips

1. Job search and interview preparation are lifelong processes.
2. The best time to interview is when you don’t need a job.
3. Start looking for jobs 3-6 months before.
4. Build up your portfolio and publish them.
5. Get people to refer you.
6. Look up your interviewers. Review their work.
7. Have your friends to give you mock interviews.
8. Don’t pretend that you know something when you don’t.
9. Don’t criticize your previous or current employers.
10. Don’t talk about your age, marital status, religion, political affiliation.
11. Have competing offers.
12. Don’t sweat it. If you tank an interview, move on.



Chip Huyen
@chipro

I'm working on a book on machine learning interviews so I've been spending the last few months talking to companies about their hiring process for ML roles. This thread is a summary of what I've learned. It will be updated as the book progresses. (1/n)

9:46 PM · Jul 18, 2019 · Twitter Web Client

View Tweet activity

582 Retweets **2.3K** Likes

Chip Huyen @chipro · Jul 18
Replying to @chipro
The average interviewer gets very little training. You start your full-time job. You shadow a few interviews. Then you're on your own. As a result, interviews are wildly different even within the same company.

3 10 113

Chip Huyen @chipro · Jul 18
I ask interviewers to give me their most frequently asked questions, and show them questions that other interviewers ask. I notice the pattern that if an interviewer doesn't know the answer to a question, they immediately flag it down as bad.

4 15 114

Chip Huyen @chipro · Jul 18
"Explain Adam optimizer" is the most frequently asked question for machine learning roles.

9 21 258

Book info:
huyenchip.com/2019/07/21/machine-learning-interviews.html

Thank you!

chip@huyenchip.com
Twitter: @chipro