Machine Learning Interviews in 3 Weeks

From beginner to pro

Susan Shu Chang
Susanshu.com

O'REILLY®

Instructor intro

- Principal Data Scientist, Elastic (of Elasticsearch)
- Previously ML in Telecom, Fintech, Social, etc.
- Speaker, incl. PyCon US, India, Canada…
- Founder, Quill Game Studios

More career guides: Susanshu.com

Week 1:

Introduction to Machine Learning roles and required skill sets

Week 1 learning objectives

You'll learn…

- The various types of machine learning roles, such as machine learning engineer, applied scientist, and more.
- How to acquire the skill sets required for each type of machine learning role.
- How to assess your interests and your current skills before deciding which machine learning role(s) to apply for.
- How to evaluate your current skills to close any gaps that may prevent you from succeeding in the interview process.

Week 1 discussion:

What stage are you at in your machine learning job search?

What’s your biggest goal?

Testing the waters?

Veteran of several interviews?

Just started applying?

Week 1 part 1:

Understand the ML field, and ideal candidate types for each role

Machine learning career

- What products do you want to build?
- What impact do you want to have?
- What’s the salary?

Confusing job titles; big mistakes

- A candidate with good skills might be applying to the “wrong roles” with the same job title.
- After looking at the job description carefully, I helped them identify better roles to apply to.

Data roles

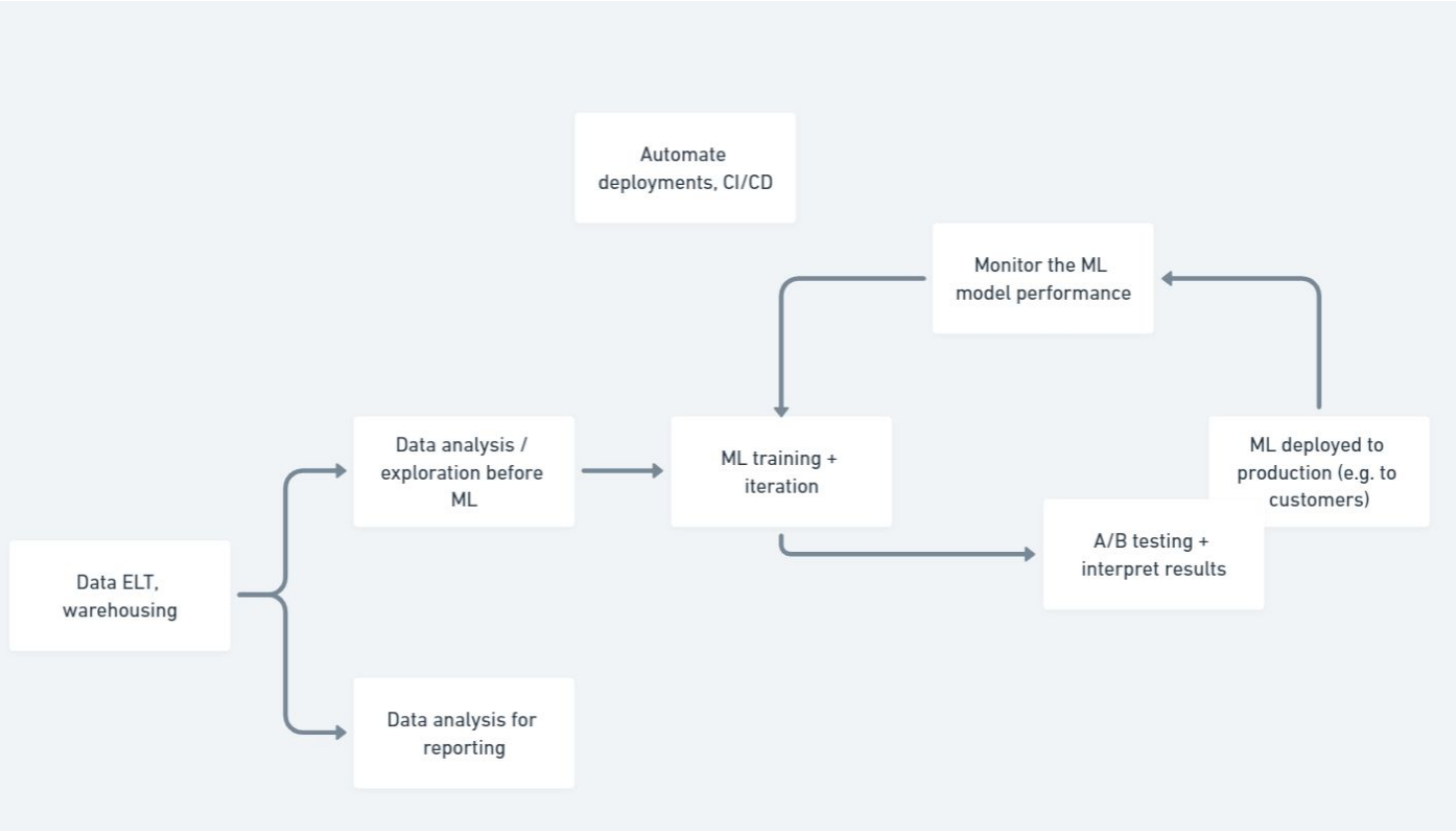
- Data Scientist
  - “Product Data Scientist”
- Machine learning engineer
- MLOps engineer
- Applied Scientist
- Data Engineer
- Data Analyst

Data roles

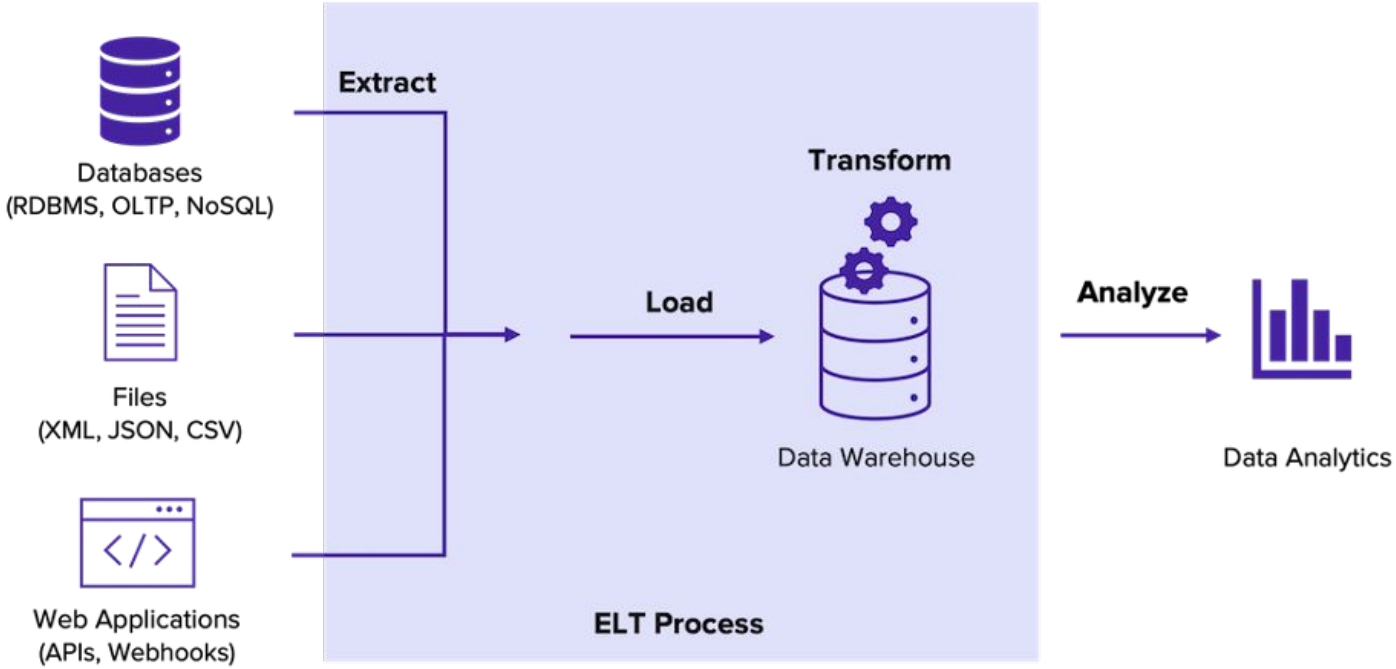
- Data Scientist
  - “Product Data Scientist”
- Machine learning engineer
- MLOps engineer
- Applied Scientist
- Data Engineer
- Data Analyst

Role confusion is a common frustration for jobseekers

The Machine Learning lifecycle



Data Engineer



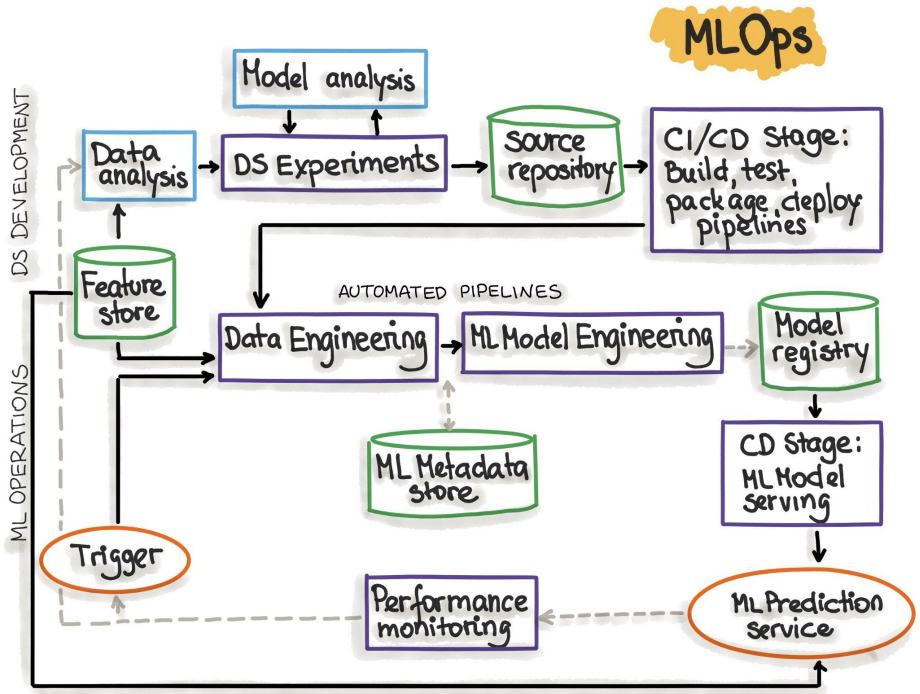
Source: [Malsha Ranawaka]

Data Scientist

- Started out as the umbrella term
- New offshoot term: “Product Data Scientist”
- Now, this job title could mean many things
    - Depends on the job description

Machine Learning Engineer

- Infrastructure work
- Model deployment, MLOps
- Model training and development
    - Part of current (newer) trend



Source: ml-ops.org

Applied Scientist

- ML models

Applied Scientists are specialists with deep expertise in at least one scientific discipline that applies to Amazon's business challenges. They also have a broad knowledge in data structures and algorithms and understand algorithmic and implementation tradeoffs.

As an Applied Scientist working in Operations Technology, you will:

- Analyze complex datasets used to make decisions regarding real-world applications
- Drive the design, development, and execution of scientific research projects and data models
- Implement models and algorithms to improve the performance of existing systems, processes, and products
- Perform data analysis and provide visualizations
- Use modeling, simulation, and optimization to design processes, algorithms, and prediction/classification systems
- Investigate the applicability of new approaches to business problems and products by combining theory and experimentation

Source: Amazon

MLOps Engineer

- Infrastructure work
- Model deployment, MLOps
- Automating ML deployments

More info on my O’Reilly MLOps Superstream keynote

Data Analyst

- Important to business decisions
  - Example: finding average sales of N business line across past X weeks
  - Automating the above via dashboards
- “Product Data Scientist”
- A good entry point for a data career

Previous trend of job titles

Data scientists do everything

Data analysts specifically responsible for data analysis related to business decisions

Current trend of job titles

Train ML models

- Data Scientist
- Machine Learning Engineer
- Applied scientist

Data analysis, A/B testing

- Data analyst
- (Product) Data Scientist

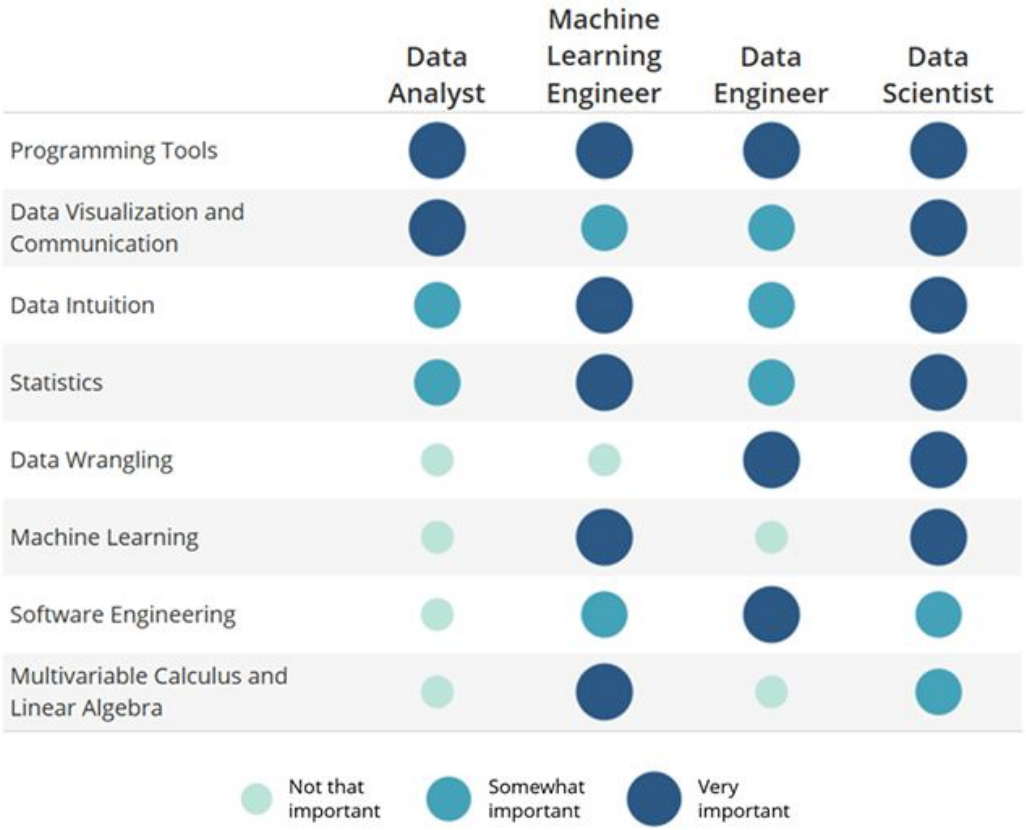
MLOps and infrastructure work

- Machine Learning Engineer
- MLOps Engineer
- Infrastructure Software Engineer, ML

Data Engineering

- Data Engineer
- Data Scientist in a startup 😛

Skills matrix



Source: [Udacity]

Week 1 part 2:

Evaluate your unique skills and create learning plan to close the skills gap

Goal of interview

Show employer that you can *do the job*
(or learn how to do it quickly)

Position yourself as **an ideal candidate**

Knowing the various ML roles, you must reflect on your unique skills.

Write down your skills.

What’s the best side project to gain the broadest set of skills?

Activity: Write down your skills!

1: Beginner
2: Experienced
3: Very experienced

Skill	Self identified skill level
Programming tools	
Data visualization and communication	
Data intuition	
Statistics	

Skill	Self identified skill level
Data wrangling	
Machine learning (theory, algorithms)	
Software engineering	
Multivariable calculus and linear algebra	

NOT the goal of interview

- The one that did a lot of LeetCode
- The one with highest scores at school

How do interviewers connect what you're saying with "the one that can do the job or learn it quickly"?

Example 1

Senior Data Scientist – Machine Learning (Social)

Bungie · United States (Remote) 1 week ago · 53 applicants

Full-time · Mid-Senior level

1,001-5,000 employees · Entertainment Providers

2 school alumni

See how you compare to 53 applicants. Retry Premium Free

Actively recruiting

Apply Save

REQUIRED SKILLS

- Experience working within cloud-based infrastructure to build, deploy, and manage models at scale in a live service environment
- Advanced proficiency in Python and SQL
- Advanced knowledge of machine learning frameworks such as scikit-learn, PyTorch, and Tensorflow
- Experience creating and using advanced machine learning algorithms and statistical models: regression, simulation, scenario analysis, clustering, decision trees, neural networks, etc.
- Effective at data visualization and storytelling geared towards both technical and non-technical audiences
- Familiarity with graph data science algorithms and graph databases (Neo4j, AWS Neptune)

Nice-to-have Skills

- Advanced degree in Computer Science, Statistics, Machine Learning, or a related field
- Experience working within a variety of different technical environments
- Experience with Tableau or similar data visualization tool
- A love of games

Bungie is seeking a Senior Data Scientist to join our newly formed Machine Learning team. Reporting to the Director of Machine Learning, you will scope, prototype, and deploy machine learning models that will support social initiatives across our organization. In this role, you will directly support Bungie social teams by building knowledge graphs in order to extract insights that drive the future of co-experience in our games.

As part of the Bungie Analytics & Insights team, you will join a diverse team of analysts, data scientists, data engineers, and researchers. You will leverage your expertise in data analysis, data science, and engineering to build machine learning services for our games and create internal tools that empower Bungie teams to continue doing what they do best. Most importantly, your work will be essential to creating unmatched entertainment experiences designed to inspire friendships and lifelong memories.

Example 2

Data Scientist or Senior Data Scientist - Core ...
Spotify · Toronto, ON (Remote)

Apply | Save | •••

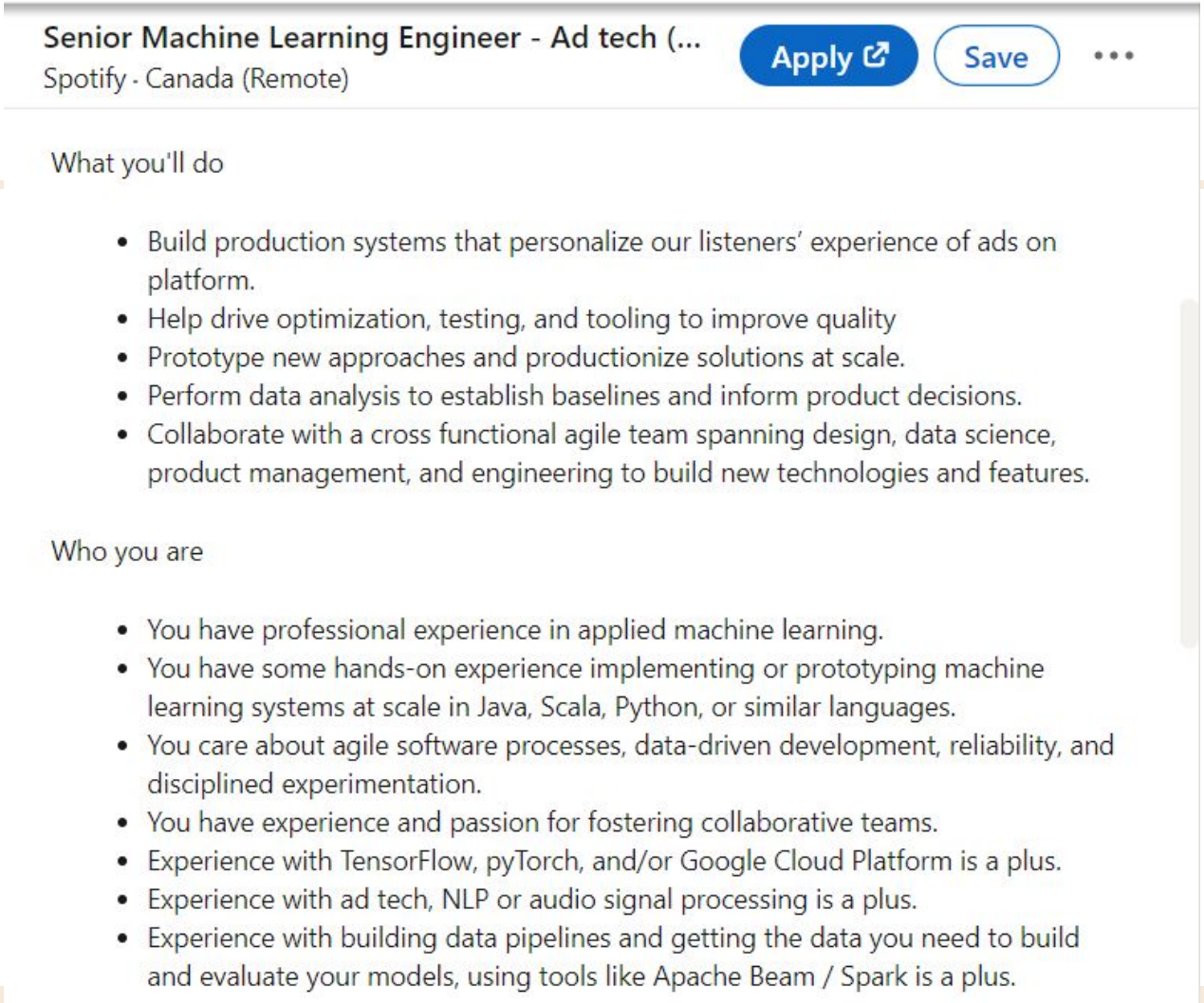
What you'll do

- Co-operate with cross-functional teams of data scientists, user researchers, product managers, designers and engineers who are passionate about our consumer experience.
- Perform analysis on large sets of data to extract impactful insights on user behaviour that will help drive product and design decisions.
- Communicate insights and recommendations to stakeholders across Spotify.
- Be a key partner in our work to build out our product strategy so that we are relevant in the daily lives of consumers.

Who you are

- Relevant experience or a degree in statistics, mathematics, computer science, engineering, economics or another quantitative subject area.
- Strong interpersonal skills and are comfortable working with multiple stakeholders.
- Know how to understand and tackle loosely defined problems and come up with relevant answers and impactful insights.
- You enjoy sharpening questions and developing hypothesis for yourself and can think together with others while clarifying their assumptions and hypothesis.
- Proficiency with Python, or similar programming languages.
- Experience with Google BigQuery or proficiency in SQL.
- Extensive experience using various analysis techniques, such as linear and logistic regression, significance testing, and statistical modeling.
- A/B testing methodologies.

Example 3



Senior Machine Learning Engineer - Ad tech (...

Spotify · Canada (Remote)

Apply Save

What you'll do

- Build production systems that personalize our listeners' experience of ads on platform.
- Help drive optimization, testing, and tooling to improve quality
- Prototype new approaches and productionize solutions at scale.
- Perform data analysis to establish baselines and inform product decisions.
- Collaborate with a cross functional agile team spanning design, data science, product management, and engineering to build new technologies and features.

Who you are

- You have professional experience in applied machine learning.
- You have some hands-on experience implementing or prototyping machine learning systems at scale in Java, Scala, Python, or similar languages.
- You care about agile software processes, data-driven development, reliability, and disciplined experimentation.
- You have experience and passion for fostering collaborative teams.
- Experience with TensorFlow, pyTorch, and/or Google Cloud Platform is a plus.
- Experience with ad tech, NLP or audio signal processing is a plus.
- Experience with building data pipelines and getting the data you need to build and evaluate your models, using tools like Apache Beam / Spark is a plus.

Common additional requirements

- git, version control
- Cross team collaboration
- Writing tests
- (Senior) Ability to break down tasks into smaller ones

Commonly mentioned on JDs, but easy to overlook.

Data Scientist

- Look at the job posting - this role varies a good amount.

You can succeed if you've:

- Trained ML, statistical models (my background)
- Also know Python well
- [Applied Scientist, MLE (depends on posting)]

Machine learning engineer (MLE)

You can succeed if:

- Trained ML, statistical models (my background)
- Also know Python well
- (Depends on posting) Worked with Dockerization, Kubernetes, CI/CD

What’s the gap between you and your target role?

How do you know there’s a gap? (1/2)

- Didn’t pass interviews
  - Was there feedback? (Companies don’t give feedback much these days)
  - What’s your self assessment?
  - Even the **job titles and descriptions** themselves can be telling

How do you know there’s a gap? (2/2)

- Didn’t pass interviews
- That’s why we needed to do the previous section
  - You could be applying for something that’s the same title, but not a match for your experience!
  - Not your fault, but now you know!
  - This training will help you identify and target well in interviews.

How to close the gap between your skills and ML role

Example goals:

- Take ~3 courses
- Build a side project yourself
- Bootcamps… yay or nay?

The BEST side project to build ML expertise

End to end ML app deployment

- Gather data via scraping etc.
- Train model and iterate
- Deploy in an app
    - Bonus: use Docker

Can 100% answer resume deep dive questions and *impress* interviewers

What’s your curriculum?

- Spend some time to write down:
- 3 Steps to take in the next 6 months
- In 3 months, I will complete…
- In 4 months, I will apply to 5 jobs a day…
- In 5 months, I will reach out to N hiring managers a week…

Next week, we'll go through…

Machine learning interview steps

- Application, resume screening
- Recruiter call (behavioral)
- Technical Interviews
  - ML theory
  - Programming
  - Case study (how would you build…)
- Behavioral interviews, cont.

“Homework”

- Look at the skills you’ve filled out this session
- Identify the gaps
- Create a 3 month / 6 month plan

Take a picture of this slide and the next!

Activity: Write down your skills!

1: Beginner
2: Experienced
3: Very experienced

Skill	Self identified skill level
Programming tools	
Data visualization and communication	
Data intuition	
Statistics	

Skill	Self identified skill level
Data wrangling	
Machine learning (theory, algorithms)	
Software engineering	
Multivariable calculus and linear algebra	

Week 1 learning outcomes

This week, you've learned…

- About the various types of machine learning roles, such as machine learning engineer, applied scientist, and more.
- How to acquire the skill sets required for each type of machine learning role.
- How to assess your interests and your current skills before deciding which machine learning role(s) to apply for.
- How to evaluate your current skills to close any gaps that may prevent you from succeeding in the interview process.

Q&A

Connect with me!

1. Newsletter with ML career tips: www.susanshu.com
2. Connect on LinkedIn: https://www.linkedin.com/in/susan-shu-chang/
3. Recommend this training to your friends! (April 2023)
4. Stay tuned for upcoming book on machine learning interviews (O'Reilly)