

ML Interview Workbook

A linked route guide and practice record for your interview loop.

Use this to prepare for the interview you actually have. Write down the loop, choose the most important gap, attempt it closed-book, record what happened, and schedule the next attempt. The workbook covers planning, practice, simulations, stories, and the final week.

Candidate: _____

Loop date: _____

How to use it

1. **Start with recent practice.** Record confirmed rounds, time available, and what you can do without notes.
2. **Choose three gaps.** Work on the areas most likely to affect your actual loop.
3. **Attempt, compare, retry.** Keep notes closed until the attempt ends; compare your answer with the criteria; schedule the next attempt.
4. **Call a skill ready carefully.** Look for two successful attempts on different days and one unfamiliar follow-up.

1. Loop snapshot

Target role: _____

Level: _____

Primary domain: _____

Weeks available: _____

Hours per week: _____

Interview date: _____

Confirmed round formats

- | | |
|---|---|
| ☐ ML breadth | ☐ ML implementation / coding |
| ☐ Agentic codebase | ☐ ML system design |
| ☐ Project deep-dive | ☐ Technical strategy |
| ☐ Technical presentation | ☐ Behavioral / leadership |
| ☐ Values & mission | ☐ Product / experimentation |
| ☐ Research depth / work sample | ☐ Math / statistics oral |
| ☐ Systems / infrastructure | ☐ Other: _____ |

For every checked round, confirm the task or prompt: _____

Duration: _____

Tools, coding environment, or presentation format: _____

Your route through the field guide

This PDF is a route guide, not your saved plan. Open [Build my plan](#), enter your confirmed rounds, role, level, domain, available weeks, and hours. The [browser Workbook](#) selects diagnostic and transfer questions, links the repair material, and tracks retry dates. Use **Print my study map** for your own routes and progress.

Choose the scope before the reading

1. Ask the recruiter for each round's task, duration, tools, and expected depth. A title alone does not specify the loop. Mark unconfirmed formats as questions to resolve.
2. Attempt one representative prompt per confirmed round. Start with the weakest evidence, not your favorite subject. A score from reading is not a baseline.
3. Reserve half your weekly time for retries, repairs, and feedback. A 40-minute design attempt plus 15-minute review uses 55 minutes, not an entire evening.
4. Add breadth only after the diagnostic identifies a missing mechanism. Two mapped prompts sample a round; they do not prove complete coverage.

Use the same method on every route

1. **Attempt.** Open the diagnostic in Practice Mode and leave the answer closed. Use its timer and round-specific rubric.
2. **Score.** Weak: cannot complete the core task. Review: workable parts, but an important gap remains. Confident: a clean timed answer with no weak rubric item. Record the exact failed step.
3. **Repair, at most 15 minutes.** Read the relevant visual or reference section. Close it and reproduce only the missing derivation, diagram, test, or decision. If it still fails, split the repair into a smaller task.
4. **Repeat later.** Weak: two days. Review or first clean Confident: seven days. Keep the saved date even when the answer now feels familiar. These are practice defaults, not guaranteed optimal intervals.
5. **Transfer.** After a clean diagnostic, try the second linked prompt without notes. After two clean passes on different days, use a changed assumption in a later mixed session. Mark the pass only after doing it.

When time runs short: reduce new topics, not the evidence standard. An unattempted or unrepaired round remains a risk. Ask a peer to probe that round before polishing already-strong answers.

Worked routes: select, attempt, repair, transfer

These examples show how to choose from the library. Use only the rounds in your actual loop; the browser map makes these choices from your inputs.

Applied Scientist: product ML, L5

Confirmed rounds: breadth, system design, project deep-dive. Eight hours per week; four weeks.

- **Design diagnostic (40 min + 15 review):** [fraud detection](#). State the decision, error costs, label delay, evaluation split, serving path, and rollback. If threshold choice is hand-waving, repair with [decision thresholds and asymmetric costs](#). Transfer: [an AI coding product](#) with a different error cost and feedback loop.
- **Breadth diagnostic:** [bias and variance](#). Explain a learning curve without reciting definitions. Repair the [estimator visual](#); transfer to [loss selection](#).
- **Project diagnostic:** [your ambitious project](#). Use one real decision, rejected alternative, measured outcome, and failure. Transfer to [an ambiguous scope](#). Do not add coding unless it is confirmed.

ML Engineer: LLM serving, L5

Confirmed rounds: implementation and system design. Five hours per week; two weeks.

- **Coding (40 + 15 min):** [implement attention](#). Test shapes, masking, numerical stability, and gradients. Repair the [attention visual](#); transfer to [KV-cache decoding](#).
- **Design (40 + 15 min):** [production LLM inference](#). Quantify prefill, decode, memory, queueing, and degradation. Repair [prefill versus decode](#); transfer to [an AI coding product](#).
- **Week budget:** 110 minutes for the two diagnostics; reserve 150 for retries and repairs; 40 remain. If Tuesday's coding attempt is Weak, retry Thursday. If Thursday is clean, repeat the following Thursday. Do not mark it ready on Friday from familiarity.

Research Scientist: paper defense and math

[Critique a paper](#) for 25 minutes: separate the claim from the evidence and strongest alternative explanation. Repair [evaluation validity](#); transfer to [an ablation](#). For a confirmed math oral, use a [15-minute derivation](#); transfer to [logistic regression](#). Add the [45-minute presentation](#) only when that format is confirmed.

Other confirmed rounds: precise entry points

Use the linked question's timer and scorecard. The same two/seven-day retry rule applies; the task changes by round.

Round	Diagnostic and observable output
Agentic codebase	Extend an evaluation codebase : a bounded change, inspected diff, and passing tests. Use the executable labs for the work environment.
Technical strategy	Enterprise agent platform : ownership boundaries, investment order, migration, stop conditions. Compare with the annotated mock .
Presentation	Technical project : claims, evidence, personal decisions, interruption, technical descent.
Behavioral	Disagreement : a real trade-off, your action, result, and update. Draft facts in the story bank , not a memorized script.
Values and mission	Values under pressure : competing obligations, cost of your decision, and evidence that changes your view.
Experimentation	Design an A/B test : unit, exposure, primary metric, guardrails, power, threats, ship rule.
Research work sample	Black-box investigation : competing hypotheses, discriminating probes, measured readout.
Infrastructure	Train a 100B model : resource estimate, bottleneck, parallelism, checkpoint and recovery plan.

2. Evidence scorecard

Score only recent timed, closed-book practice. **0** means not attempted; **3** means you can give a workable timed answer; **5** means two successful attempts on different days plus an unfamiliar follow-up.

Area	0	1	2	3	4	5 / note
Fundamentals	☐	☐	☐	☐	☐	_____
Production / systems	☐	☐	☐	☐	☐	_____
LLM systems	☐	☐	☐	☐	☐	_____
Retrieval / search	☐	☐	☐	☐	☐	_____
System design	☐	☐	☐	☐	☐	_____
Experimentation	☐	☐	☐	☐	☐	_____
Behavioral / leadership	☐	☐	☐	☐	☐	_____
Math / research	☐	☐	☐	☐	☐	_____
Coding / implementation	☐	☐	☐	☐	☐	_____

3. First three practice priorities

Choose gaps by importance and evidence. Each plan item should name an action, not just a topic.

01 Gap / round: _____
 Closed-book action: _____
 What success looks like: _____
 Retry date: _____

02 Gap / round: _____
 Closed-book action: _____
 What success looks like: _____
 Retry date: _____

03 Gap / round: _____
 Closed-book action: _____
 What success looks like: _____
 Retry date: _____

4. Story bank prompts

For each story, keep the ownership and evidence explicit. **Problem / stakes**

My decision

Alternatives rejected

Evidence / result

Failure or surprise

What I would change

5. Weekly operating loop

Start of week	Check due retries first. Reserve half the weekly budget for repair and feedback. Name one new diagnostic only if time remains.
Practice days	Attempt closed-book. Use a visual lesson only after the attempt. Record the failed step and saved retry date.
On the due date	Weak: two days later. Review or first clean Confident: seven days later. Do not force all retries onto Friday.
Weekend	Review the evidence, update the plan, and remove work that is no longer relevant.
This week's focus	

What changed after an attempt?

Next action

6. Readiness check

Do not call a critical gap ready based on confidence alone.

- Two successful closed-book attempts on different days.
- No critical scoring area remains weak.
- One unfamiliar follow-up or changed assumption goes well.
- One later mixed practice session remains workable.
- A peer, mentor, or interviewer has challenged at least one claim when possible.

What I still need to show:

Use one row per meaningful attempt. Record what you could do without notes and what changed after feedback.

Retry protocol

Before the attempt	What must I do without notes? _____
After the attempt	What could I not do yet? _____
Next practice	What is the smallest task that addresses that gap? _____
Evidence	What should improve next time? _____

8. Simulation review

Use this after a role-specific simulation or a full practice round. Separate a smooth answer from evidence that holds when an assumption changes. **Simulation / role:** _____ **Date:** _____

Rounds included: _____

Time limit and conditions: _____

What held under pressure? _____

Where did I lose structure? _____

Which assumption changed the answer? _____

What evidence was missing? _____

What will I practice before the next simulation? _____

Observer / interviewer feedback

Strongest signal

Most important improvement

Follow-up I should rehearse

9. Project and story practice

Keep the story focused on decisions, ownership, evidence, failure, and what you changed. Do not memorize a script.

90-second opening Problem, ownership, outcome, and the three decisions I will explain.

Decision 1 Constraint: _____

Choice and evidence:

Decision 2 Constraint: _____

Choice and evidence:

Decision 3 Constraint: _____

Choice and evidence:

Ownership What did I personally decide, build, unblock, or change?

Failure / surprise What failed, and how did it update the plan?

Impact and guardrail What changed for users or the business, and what could have regressed?

Changed-assumption follow-ups

01. If the data, latency, budget, or label changes:

02. If the first approach fails:

03. If a senior partner disagrees:

10. Other interview rounds

Use this page for formats outside the guide. Those formats need their own practice plan.

External format	Separate resource or drill	Baseline date
General algorithms / DSA	_____	
Practical software / OOP / concurrency	_____	
SQL / data manipulation	_____	
General distributed systems	_____	
Other:	_____	

11. Final week

The final week is for steady practice on known gaps, not broad new coverage.

- ☐ Recheck each critical round with one closed-book baseline.
- ☐ Retry only the most important weak areas.
- ☐ Run one mixed simulation with an unfamiliar follow-up.
- ☐ Practice the 90-second project opening and three decisions.
- ☐ Confirm time zone, links, format, environment, and permitted tools.
- ☐ Prepare concise questions about scope, success, collaboration, and the first 90 days.
- ☐ Stop adding broad topics; protect sleep and response quality.

Logistics / questions / risks

12. Review and backup

At the end of each week, update the plan from what you practiced, not from how you felt.

What did I complete?

What remains critical?

What should I remove or defer?

What is next week's one priority?

Backup reminder

The browser workbook is local to the device. Export its JSON backup before changing devices or clearing browser data. This printable PDF contains no private plan data.

A compact operating rule

Practice without notes. Compare your answer with the criteria. Work on one gap. Retry after time has passed. Study a topic only when it changes what you can do in the interview.