CV Reengineering Leader Interview Protocol

Interview Questions

This interview is related to the re-engineering process of DL models.

We have five sets of questions here: Background, Process Workflow, Challenges, Effective Practices, and Others.

Background

First, I would like to ask some warm-up questions for our interview, in order to know more about your background in DL/SE. The background questions can be answered briefly so that we can save time for the other interesting questions.

1. Could you please estimate your {DL, SE} expertise level? (Novice, Intermediate, Expert, Master.)

2. How many years have you worked on {DL, SE} professionally?

3. What is your educational background?

    a)   Where did you learn about {DL, SE}?

4. Roughly, how many {DL, SE} projects have you worked on?

    5.   Do you work in a team or individually?

        a)   What is the size of your company? (Small, medium, large.)

        b)   How many people are there in your team?

        c)   What is your role in your team? (Manager, leader, contributor)

    6.   What kinds of models have your worked on? (E.g., CV, NLP, Game)

7. Think about the projects where you/your team re-engineer a DL model using research papers/existing implementation?

    a)   What are your goals of re-engineer such model?

        i.   Why are you doing this?

    b)   *Did you find that the context was similar enough for the work to be directly applicable? How much did you adapt the model, loss function, etc. to make it work in your context?

Process

This next set of questions is related to the DL re-engineering process. We are trying to understand the process that software engineers follows as they try to bring DL research

results into their organization.

1. Can you talk me through the process that your team follows to re-engineer a machine learning model from research paper/existing implementation/another engineer's project?

    a. How do you … [reading the paper, implementing, unit testing, optimizing, debugging, training efficiently, etc.]

    b. (Show the process workflow) Then I will show you a process workflow. Please take a look at the workflow. I will pause the record for a minute, and then I will ask several questions. You can tell me when you are ready.

    c. Can you tell me if you think this is an **accurate** process workflow, or if you would change things based on your team's actual process?

        i. How do you measure success of your re-engineering?

        ii. The workflow does not contain a "**Feature Engineering**" stage. Does your re-engineering work involve such a stage? If so, where should we add this stage in the diagram.

2. *Would you like to add any back-edges in this diagram?

    a. *How does your team update new iterations of your model if it doesn't work for the first time?

Challenges

The next set of questions is related to challenges in DL re-engineering process. We are trying to find hot spots in this process.

1. Are there any hot spots in the DL re-engineering process or which parts do you think are challenging when re-engineering a model, e.g., understanding the paper, data pipeline model, loss function, optimization, etc.?

    a. Can you tell me about an error you found in a **LOSS FUNCTION/MODEL/DATA PIPELINE/Evaluation/Optimization?** (The error can be traditional bugs, bad performance, memory exhaustion, etc.)

2. Can you describe any challenges you met when implementing a LOSS FUNCTION/MODEL/DATA PIPELINE/Evaluation/Optimization/Documentation?

    a. How do you address these challenges?

    b. *What kind of documentation would be helpful, BEYOND paper + source code?

    c. *What do you mean by "…"?

    d. *What difficulties arise here?

3. *How much of an impact can a [*bug they mentioned in previous question*] have in different stages of the DL re-engineering process?

    a. Have you met any challenges when integrating all components?

    b. *How much does it affect your team in the implementing/optimization stage? (Does it stop your process?)

4. Based on your experiences, I wonder if you can think of 1 – 2 changes to

processes/tools/artifacts that would make this process easier for you. Please think for a moment and then let me know your thoughts.

Effective Practices

The next set of questions is on effective re-engineering practices. We would like to know what practice you and your team adopt in your engineering process (making the re-implementation/cooperation more effective/efficiently).

1. Have you ever reused any pre-trained models?
   a. How do you assess the viability for reusing pre-trained models?
   b. *Can you give any examples of successful/failure reuse?
2. How does your **team** work together to make the process more effective?
3. How do you decide an **existing implementation** is trustworthy?
   a. What could **other** engineers (who contribute to open-source projects, group in company, scientists) do differently in their implementations? Why? *(If they mentioned this.)*
4. What do you find is helpful/problematic in a **DL research paper**?
   a. What could DL **researchers** do differently in their **papers/implementations** to help you re-engineer the work? Why?
5. What do you find is helpful/problematic in the **documentation** of DL models**?**
   a. What could documentations be different to help you re-engineer the work? Why?
6. Are there existing tools (or other technologies) you found valuable/problematic for re-engineering, e.g., ONNX (Open Neural Network Exchange), testing tools, visualization tools?
7. How do you determine the acceptable trade-off between the performance of the model (accuracy/speed) and the cost of your team (time/money, etc.)?

Other

(The last set of questions is comparison between DL re-engineering process and other SE process. We are trying to find the specific feature of DL re-engineering process.)

1. (The next questions are related to **Scientific & Engineering Effort** in this process.)

2. *In what ways do you think the re-engineering process is a scientific effort? And in what way it's an engineering effort?

3. (You mentioned you have experience on general DL process.) What do you think are different aspects of DL re-engineering compared to a general DL developing process?

4. (You mentioned you have experience on SE process.) Can you talk about the difference between DL re-engineering process and other SE process?

5. What recommendations/suggestions would you like to provide to engineers
   i. Who are experts in SE but novice in DL if they are involved in DL re-engineering process?
   ii. Who are experts in DL but novice in SE?

*Are there any other questions you would like to know/ask other engineers which is related to such re-engineering process?

Anonymization

Audio from the interviews will be recorded and transcribed by a trusted 3rd party service. The transcriptions will be viewed, and personal information will be anonymized. The audio recordings will be deleted after the transcripts are received, reviewed, and anonymized. A key mapping participant to the transcript will be kept separately by the research team until the conclusion of the study – this key may inform further participant selection, e.g. further probing of an interesting organization.