Interview 13 - Rached

Interviewer 1

There. We all right. Okay. So, uh, we'll start by introducing ourselves and me. It's, you maybe already saw me. Mm-hmm. in labs. Um, I'm Master Sudan and I'm under the supervision of, uh, foot and I'm in, was not there today. Asha, if you want to present yourself.

Interviewee

I'm, I'm, and I'm doing my master research and, uh, I'm also on the of profess.

Interviewee

And, uh, yeah, that's it. Thank you for, for coming to Pleasure.

Interviewer 1

Yeah. The, were you about, if you want, you can present yourself. It's.

Interviewee

Okay, so, hi, my name is X. I'm a data scientist at Company 1, um, for about two years now. And, uh, yeah, that's it. Okay, perfect.

Interviewer 1

Thank you. Uh, so yeah, so the goal of the interview is to, in the end goal is to develop a catalog of quality issues in machine learning software system.

Interviewer 1

And what is the machine learning software system? It is just any software system, but that has a machine learning component in it. Okay. Yeah, not much more complicated than that. And quality issues. Uh, yeah. So any issue with the quality, any issue with the quality of ML system. Okay.

Interviewee

Yeah.

Interviewer 1

Um, So, yeah, we'll ask, uh, you, I will start the interview.

Interviewer 1

We have about like 20, 25 question. Okay. And we are not expecting you to know, like, to have experience in on every question we ask you because sometimes it's more like data scientist, sometimes it's more data engineer. Okay. So if you don't know, we'll just keep the questions. Sounds good. Perfect. All right.

Interviewer 1

Uh, so what are the main quality issues you, you have encountered, uh, with your data model or system so far? In general? In general?

Interviewee

Yeah. Um,

Interviewee

I would say that often when we build a ML system at Company 1, we, um, we will, we will work with data that is not. Very clean or very well collected, gathered or organized. So, um, it'll happen often that because of the quality of the data, uh, this will impact the quality of the, um, ML system down the road.

Interviewee

Okay. Um, that's one of the main aspects. Then I would say, um, Sometimes, um, the vulgarization of ML concept to a client or to a business stakeholder is not always straightforward. And sometimes the understanding of the stakeholder of some ML concept or statistical concept is gonna affect the quality of the ML system because, um, Expectation will not be met.

Interviewee

Uh, what we will deliver might not be exactly what the client is expecting. So, so some misunderstanding, uh, usually I think affect the quality of what is delivered. And

Interviewee

yeah, I would say those are the two main points. Okay, thanks.

Interviewer 1

And so you mentioned issue with the quality of the data. Mm-hmm. , what are some of the issues you have

encountered?

Interviewee

Uh, so for example, um, Sometimes the client thinks that he records some information, but in reality, he's not

recording that information.

Interviewee

He's recording something else. So we don't have access to the information we needed to build the ml, uh,

system. Sometimes there are a lot of missing values. Sometimes they're, um, outliers. Sometimes they. Um,

discrepancies in the data, um, not enough data that's, that has happened. Often that decline thinks that he

collected a lot of data, and that's enough to build an ML system, but in reality there's, uh, very few data points.

Interviewee

So

Interviewer 1

those are some examples. I see. Thanks. Um, did you ever use, To, to gather data, to train your model. Did

you ever use, um, like manual labeler, people that collect

Interviewee

data for you? Uh, never. No.

Interviewer 1

Never. Okay. Uh, did you ever use external data? So, uh, for example, public data set or third party API or

web script data, something like that?

Interviewer 1

Yeah.

Interviewee

That has happened more often.

Interviewer 1

Yeah. Yeah. And what are the, the quality you have with the data when you use these type of. Data collection

Interviewee

technique. Uh, usually when we use a third party provider, the quality of the data is uh, very, uh, mu uh, it's much better. Uh, we usually don't encounter too much issues, so I would say I see a very big difference between third party providers and, for example, client data.

Interviewer 1

Okay. I see. Oh, that's interesting. And did you ever use data that was generated by another system? Um, so and yeah.

Go

Interviewee

for it for another, by another system. Like for example, uh,

Interviewer 1

yeah, by example. Um, I mean, it's a bit vague. Um, for example, um, your client could have, uh, some system that produce data mm-hmm. , and you use it in your ML models or in your data pipelines.

Interviewer 1

For the model at the end, has it ever happened to you?

Interviewee

Yeah, so usually the client has some machines or a transactional system or something like that, that collects data every day. And we don't, uh, plug into those system directly, but we plug into the data warehouse that already has processed the data. So, um, so usually the data is already somewhat processed and clean.

Interviewer 1

Okay. So no issue with that? No. Okay, thanks. Uh, so moving on to data preparation. Um, have you ever measured the quality of your data and or tried to improve it?

Interviewee

Uh, yeah. Usually, usually when we build a ML system, depending on the use case, we will run some test or analysis to verify that the, that the data is, um, uh, usable for the goal we we wanna achieve.

Interviewee

And if not, we. We will, uh, do some transformations or some cleaning or some kind of data wrangling to make it, uh, in the right format or in the right shape to be able to use it.

Interviewer 1

Okay, I see. And how do you do it and which tool do you use to do it?

Interviewee

Um, we usually do it by, by hand, , meaning we analyze it, the data with like, for, for example, a Jupyter Notebook doing some, uh, statistical descriptive statistics, for example.

Interviewee

Um, and, um, And when we encounter a problem, that's where we kind of think, okay, how could we solve that, that problem? For example, and I can give you an example where we were trying to do forecast time series forecasting, and a lot of time series were very, uh, sparse and so it was difficult to model those time series.

Interviewee

So in, in order to be able. Build an ML system with those time series we a it some time series that were related to make them more smooth

Interviewer 1

modeling friendly. Okay. Okay. Thank you. Um, and are the, are they do, is there some issue? You encountered a lot of time with the data that almost every day you have this.

Interviewee

Um,

Interviewee

I'd say that almost every time we will encounter outliers, we'll encounter, um, values that are impossible. For example, negative sales, for example, our clients. So, um, um, we will always encounter as well. Uh, text pills that are not clean or not standardized, which we have to clean. Uh, things like that.

Interviewer 1

Okay, perfect.

Interviewer 1

Thanks. And is there any other data quality issue we miss that you consider relevant? Um,

Interviewee

uh, not on the top of my head. No. Perfect.

Interviewer 1

So now we've been moving on to modeling question. If you, if you feel comfortable

Interviewee

or not, tell us. Mm-hmm. ,

Interviewer 1

right. So do you know how the, uh, that data scientists evaluate the quality of their models? Um,

Interviewee

mainly through somes or, uh, performance metrics? Um, yeah, mainly with that.

Interviewer 1

Okay. Perfect. Thanks. Um,

Interviewer 1

do, do they assess a quality of, uh, ML model prediction with the user of the system? Um,

Interviewee

that's a good question. Sometimes yes, sometimes no. And by our experience, when we didn't assess the quality of the model with the end user, that resulted in a mismatch of understanding and expectation. So now it's becoming more and more and, uh, uh, mandatory to have some kind.

Interviewee

Evaluation with the client or with the user to make sure that at the end when it's, when he's gonna use it, it's gonna give the result that we are intended.

Interviewer 1

I see. And how do they proceed when they're with the client? Do they just present randomly some predictions or they have some way of presenting the ML system?

Interviewee

Usually it's to a. Like a project where, uh, each day, for example, we will generate some live predictions. We will send them to the user for review, and then he will give some feedback about those predictions or those, the outcome of the ML system and, um, uh, thinking and account. Those feedback. We will, uh, for example, either explain the result to him or we.

Interviewee

Fix something in the model and further down the road we will then compare the, uh, output of the ML system with the actual value that was expected. And then, uh, have another feedback session with the end user to make sure that there was no discrepancies are. Okay. I see.

Interviewer 1

Interesting. Thank you. Uh, now moving on to model Deploy.

Interviewer 1

Um, how and where are your models deployed?

Interviewee

Um, almost always using a pla, a cloud platform. Okay. Yeah.

Interviewer 1

Is it generally, uh, do you use all three cloud platform or you have one

Interviewee

preferred? Uh, we use the cloud platform that is used by the, the client or. The person we are working with. So we, so far I've used both, all, all three of them.

Interviewee

All the three, the main ones. Okay, I see.

Interviewer 1

Um, what are the challenge that you have encountered during the deployment of machine versus, or the main challenge you have? Usually?

Interviewee

Um, main challenge. Um, some, sometimes when we. We do like a proof of concept. We do a first ML system on the side that to prove a point or to prove that there's value added to go some way or to use some data to achieve a purpose.

Interviewee

Then when we, um, deploy this algorithm or this ML system, Sometimes we didn't account for some like business logic or, or business constraint that, uh, that prevents us from deploying the ML system as is. We usually have to, uh, um, go back to the workable and think of some changes we need to do in order to deploy it in, in the real world in.

Interviewee

The business operation, the day to day. Okay.

Interviewer 1

Interesting. Could you give me an example if it's possible with you? Uh,

Interviewee

sure. Um, for example, a silly example would be, uh, we, we needed the weather, um, weather data to, uh, uh, in a animal system to predict some kind of value for a client. And, um, During the, uh, proof of concept, we used some historical, uh, weather data that we had to test and make sure that it was doable.

Interviewee

And then, uh, when we deployed the model, we, we, uh, noticed that we couldn't have the weather data up until like, uh, today or yesterday. There was always a gap of. Three days because of the provider we were using. So because of that, we needed to adjust our ML system to make predictions further down, uh, in order to account for those days that we missed the weather data, for example.

Interviewee

Okay.

Interviewer 1

It's pretty interesting. So some, some sort of data leak through time. Exactly. Exactly. Thanks. And you also mentioned business logic. Do you have an example for that? Mm-hmm. ,

Interviewee

um, For example, uh, uh, another ML system we built was to, uh, forecast the, um, the quantity, uh, needed for some kind of, uh, food preparation of some ingredients.

Interviewee

Um, but we didn't account for the fact that, uh, sometimes, um, With one single uh, ingredient, you can do several products. Uh, so, so yeah, we forecasted each product instead of forecasting the main ingredient that was used to build the other products. I'm not sure if it's care.

Interviewer 1

The problem is a problem that you were pretty, for example, you.

Interviewer 1

You were doing a tomato soup and a tomato juice and you predict only one tomato instead of two tomato,

something. Exactly. Yeah, something like that. Okay, so I understand you, you were missing like the, uh, I

understand.

Interviewee

Yeah. The, like the food transformation process. Yeah. You cannot only predict the, the, the end product.

Interviewee

You needed to predict the, the, the raw product as. So, um, but we, we didn't account for that, so we were

predicting for everything . Okay. The tomatoes and the tomato soup. Okay.

Interviewer 1

So just to clarify, so you had a, like a set of product bought by your consumer mm-hmm. and you wanted to

predict recipe from what they bought.

Interviewer 1

Is

Interviewee

that right? So we have. I'm gonna start again. We have transform products, right? So transform products

could be a tomato soup, uh, tomato juice, whatever, . And we have raw products, which would be just the

tomato. Okay? So the goal of the, of the client was to predict how much tomato do I. To build those other

transform products.

Interviewee

Oh, okay. Okay. Um, but, um, but we didn't account for that. We predicted instead the numbers of final

transform products. Okay. Uh, and then we, we, uh, so we predicted that the number of those final products

and the, the number of the raw product as. But independently, we didn't account for the relationship between

the two.

Interviewee

Okay. Um,

Interviewer 1

so, so let me, sorry. I think it's like the turn timeline I'm trying to recollec. So, so, um, you have a bunch of, uh,

transform product and you try to predict how much raw product you need. Something like that. Mm-hmm. .

Exactly. Okay. And can you do it like with, uh, with logic instead of machinery?

Interviewee

Uh, we could have done it with logic, but we, this logic wasn't developed by client, so

Interviewer 1

Ah, okay.

Interviewer 1

It will have be troublesome to know that this product needs that something. Mm-hmm. I see's. Interesting.

Interviewee

All right. Uh, yeah. Um.

Interviewer 1

Have you encountered any other quality issue with your model or system during the deployment phase?

Interviewee

Um, uh, usually there are issues with, um, the ease of deployment of some algorithms or some libraries that.

Interviewee

For example, uh, uh, we use them in, uh, proof of we use some algorithm and then, uh, during the deployment

phase, we, uh, figure out that there's so much data that we need to parallelize this algorithm and, uh, It's not

straightforward because it's not supported by this tool or that tool that is using that club PLA platform.

Interviewee

So we have to build some custom solution and usually to be able to translate the proof concept to deployment for some parts of the solution, that can be

Interviewer 1

difficult. Okay, so sometime models are doing prediction too.

Interviewee

Uh, sometimes models, uh, are taking too much time to compute, so we, we need to parallelize them.

Interviewee

But sometimes in the, it's not straightforward because it's not supported out of the bugs by the algorithm. So we have to find some, some other ways to parallelize it, and that can be challenging.

Interviewer 1

Okay. This is for training time or for inference time.

Interviewee

Usually train time. Okay. I see.

Interviewer 1

Can you, do you have an example in your head?

Interviewee

Uh, yeah. For example, we, uh, used, uh, the, uh, light GBM algorithm for some ML project. And the one that we used, uh, couldn't be paralyzed, so we had to, um, to switch to, uh, spark ml and then parallelize all our data processing and all, all the training inference. So that was, um, major architecture change from the prove of concept to the deployment because of the limitation of I gbm outta the

Interviewer 1

Okay.

Interviewer 1

Super.

Interviewer 1

And so basically the model cannot be paralyzed. Mm-hmm. , I'm going a bit deeper. The model cannot be

paralyzed, so you, you use an elaborate that make makes it easier to. Your code. Okay. Exactly. Mm-hmm. ,

do you know if there's any, uh, like, um, model that can be paralyzed? I'm just, uh, trying something. Yeah.

Interviewee

Um. I think that most models can be implemented with parallelization, but uh, it depends which library are

available or not.

Interviewee

Okay. Which, which algorithms? A algorithms were, um, uh, build to be able to be paralyzed or not, so, yeah.

Interviewer 1

I see. And which library did you use for

Interviewee

reli? So, uh, so we, For the approval concept, we were using the like GBM library. And then for the

deployment we use synapse ML library. Okay. Which is an implementation of a grad and boosted tree.

Interviewee

Uh, but that can be paralleled.

Interviewer 1

Okay. Thanks. Uh, now moving on to maintenance. Try and you tell me if you don't know. Okay. Uh, so how

do you ensure that the quality of a machine learning system does not decrease over time?

Interviewee

Uh, usually we monitor regularly some, uh, key metrics, performance metrics we look at. We also look at data

drift, for example, to make sure that the data is still. Similar to what we used to see. Um, we, uh, always make

sure that, um, the, the system is, um, uh, is still used by the user because, uh, usually a sign that the, uh, ML

system is not useful or there's a issue is the user is not using it.

Interviewee

So we always have like, Feedback from the user to make sure it's still used and useful. So I would say that those are the three main things we did, we look at. Okay,

Interviewer 1

super. Thank you. Um, did you have ever encountered, um, yeah. Issue with the data during the maintenance of a machine learning software system?

Interviewer 1

Um,

Interviewee

I'll say not during maintenance, but, uh, but the user, um, got back to us with some problems or some issues that he was seeing in the predictions or in the output of the ML system that then we fixed and we, um, we, uh, founded the issue and we fixed it.

Interviewer 1

Okay. And the issue was it in data.

Interviewee

Uh, usually, yeah, it was in the data processing, data logic, uh, uh, that was put in place that sometime breaks.

Interviewer 1

Okay, I see. And did you, do you have mechanism, did you put some mechanism to prevent these kind of issues from happening in the future or to see them coming before it becomes an. ,

Interviewee

that's a good question. At that time, we didn't have any mechanism, but after that issue, we put in place some tools like, uh, great Expectations, which is a library that enables to verify some, some aspects of your data types, data columns, things like that to make sure that you see it way before it breaks in.

Interviewee

DMS system. So now, yes, we use some mechanisms too.

Interviewer 1

Okay. So you have great expectation. Do you have any other tool?

Interviewee

Uh, sometimes we also build, uh, custom scripts to verify particular aspects of the data. Uh, but mainly we use great expectations. Okay. Perfect. Thank you.

Interviewer 1

Uh, have you encountered issue with the.

Interviewer 1

During the maintenance of a machine learning software system?

Interviewee

Not really, no.

Interviewer 1

Okay.

Interviewer 1

Uh, do you have any other issue regarding the maintenance of, uh, machine learning software system?

Interviewee

Um, not really. No. Thanks.

Interviewer 1

Uh, so now we are moving on to the last section of the interview. Okay. So I will just ask you a question. About some quality measure in a machine learning. Yeah. Thank you. Uh, so did you ever encounter the issue?

Interviewer 1

Did you ever add issue with one of the following aspect? So either fairness or robustness, explainability, scalability, or privacy? I can go back to any of them if you want.

Interviewee

Mm. So the last two were explainability and scale. Sorry.

Interviewer 1

Scalability. Yeah. Uh, scalability and explainability. Yeah.

Interviewee

Okay. Uh, so we encountered issue in explainability and scalability in the past.

Interviewee

Yeah. Scalability, because, um, the client want to scale to, uh, new, uh, new stores or new products or new, uh, add more, uh, scope to the, the ML system. And the ML system wasn't able to scale at that point and explainability because the user was very septic about the output of the ML system and wanted to understand more, right.

Interviewee

Why is it predicting this or that?

Interviewer 1

Okay. So and you go a bit deeper in the scalability. Uh, sorry.

Interviewee

Um, so we. System for, uh, for business process for the client, but then he wanted to replicate that ML system to similar process in the same. In the same company. So, uh, but when building the ML system, we take it, we took into consideration the constraint of that business process in particular, for example, number of products,

number of data points, number of days in the years.

Interviewee

And then when he wanted to scale to similar business process and shared some curs. The ML system wasn't able to scale because there was so much more data and, uh, there was, there were ed Edge cases that the other business process were not, uh, copy paste. The first one, there was some adjustment to be made, so the ML system wasn't able to scale out of the box.

Interviewee

It needed some, uh, re-engineering. Okay,

Interviewer 1

perfect. Thanks. And for the explainability

Interviewee

mm-hmm. , would you be able to Yeah. Uh, so the explainability, um, the main issue was that the end user of the ML system was used to, uh, rely only upon the, his intuition. So usually they have a lot of, um, they have many years of experience.

Interviewee

They rely a lot of their, on their intuition to make decisions. And so, um, now that we, uh, giving them, we have given them the ML system to work with, uh, they were a bit skeptical on the quality of the output that they were. That they were given. So, um, they would always question, why is this, uh, prediction made?

Interviewee

Uh, this doesn't make sense. Or, uh, I would've taken another decision in that particular case. So the issue was to be able to, uh, gain their confide. And to be able to gain their confidence, we needed to be, uh, more transparent and be able to explain, to explain, uh, the reason why Thel system was giving that prediction or that other prediction or, um, recommending that decision versus another decision.

Interviewee

Mm-hmm. .

Interviewer 1

Okay. Thank you. And how did you do, how did you manage to explain your

Interviewee

predictions? So, uh, to be able to explain the prediction, what we did is that we chose first an algorithm that was more explainable than some type of, uh, neural network that is less explainable. And then we use some tools like the sharply values or feature importance or, uh, we also build a, a dashboard with, um, um, With, uh, prediction insights, whether why, um, uh, feature is more important than another one for that data point in particular, for example, that was interactive so the user could, um, dig deep on into the reasons why a prediction was made by DL System.

Interviewee

System.

Interviewer 1

Yeah. Okay. I see. And you, you said you use a simpler model than Dec Currenting. Mm-hmm. , did you replace your model? At the end, the product, was it the simpler model or The simpler model was only used to explain the more complex model.

Interviewee

At the end, the final product was the simpler model because it, um, it met the business needs and was less complex so, and was more explainable.

Interviewee

So we went for that solution, .

Interviewer 1

Okay, thanks. And one last question. Uh, did you, well, one before last. Uh, did you ever have issue with privacy? And your ML system?

Interviewee

Mm. Privacy?

Interviewee

Not that I can think of, no. Okay. Thanks.

Interviewer 1

All right. Um. Two quick question actually. So I, I've said last one, but it's gonna be quick. Uh, in your opinion, what is the most pressing quality issue, uh, researchers should try to solve? Or what is the most pressing quality issue in machine learning support system?

Interviewer 1

Um,

Interviewee

I would say the most pressing, pressing issue for me would be, um,

Interviewee

a way to make sure that the, that data is always, uh, is always, uh, In the, in the good format or in the, that data is always a quality before modeling or before going further. Uh, in order to avoid, avoid the looping forever between modeling and then, uh, going back to the data, uh, data processing and then going back to modeling.

Interviewee

I think that, um, Most of the issues that I've encountered comes from the data quality and data processing. So in my opinion, that would be the main challenge. Okay.

Interviewer 1

Uh, do, do you mean really like the quality of data? What will be great is it to know the level of quality of the data or to know if you can use that data to, for your.

Interviewee

Um, I think, uh, a mix of both, like, um, uh, sometimes, um, have no experience that some data scientists jump too far to, uh, to modeling and to trying to solve the task with data that can not even be used to solve the, the test. So that's one thing. But also, um, sometimes the, the daylight itself is not, is. Clean enough or is not like you could, uh, you could put in place some tools to make sure that your data get is always at a certain, uh, level of, uh, quality.

Interviewee

So yeah, both, both of the those.

Interviewer 1

Okay, I see. So knowing how bad your data is and also having a tool to make it to a level that is acceptable for modeling Exactly. Thanks. Very good. Very interesting. I mean, very good. Sorry, uh, do you have any other comment about the quality of machine learning software system?

Interviewee

Um, no, I think that's it, .

Interviewer 1

Okay, perfect. Well, thank you. Uh, no, it's finished. So I think you provided very interesting. And I think it, it, it was a tough challenge since we already passed and interview a lot of people. So and you, you managed to new things. So

Interviewee

yeah. Thank you. Japan.

Interviewer 1

all right's.

Created with the Delve Qualitative Analysis Tool (http://www.delvetool.com)