Interview 14 - Amin

Interviewer 1

Okay. All right. Uh, okay. So, uh, hi, uh, I'll present myself. So I'm td I'm a master's student at Poly Technical

under the supervision of fu and I'm is not be there today. Uh, rash. If you want to present yourself. Uh, go.

Interviewer 2

I'm, I'm rash and, uh, I'm master student. I'm also, uh, uh, I'm supervised by Professor Putz and Amin and I

work in as.

Interviewee

Okay, cool. Thanks.

Interviewer 1

Perfect. Um, yeah, go for it. Yeah, I guess I'm,

Interviewee

uh, I can present myself, so I'm a X at Company X. And, uh, yeah, I guess that's pretty much it. ?

Interviewer 1

Yes. Yes. Happy to meet you. It was really interesting. Uh, the, that talk gave, I was there didn ask question,

but I was there and I, I found it interest.

Interviewer 1

Yeah. Okay. Thanks. Yeah, yeah. All right. Um, so we give you a quick description of the goals of the

interview. Mm-hmm. , yes. So in dm, what we want to add is a, a catalog of quality issues in machine learning

software system. So machine learning software system, just a system in software engineering that has a

master learning component.

Interviewer 1

and what in the quality issue, while it's any issue that, uh, for example, given you, given two machine learning, software system, if you can say that one is better than another. Well, the one that is worse probably has some quality issues. Mm-hmm. and some example of quality aspects, uh, is robustness. Explain Excellent explainability, uh, scalability.

Interviewer 1

Mm-hmm. , uh, you named it. And during this interview we will ask around 20 or 25 question. Mm-hmm. , uh, we do not expect you to answer necessarily every question because some people have more expertise in data, some people have more expertise in modeling. Mm-hmm. . So, uh, so yeah. Alright. Sounds good. Perfect. Um, you all are, uh, yeah.

Interviewer 1

I have your permission to record the interview. Uh,

Interviewee

yeah. But, um, maybe just for your own use at the moment, just for, uh, taking notes out of it, I think, and then we'll see later. Yeah. Perfect. Perfect.

Interviewer 1

Alright. Uh, so we'll start with the first question. Um, what are the main quality issues you have encountered, uh, with your data model or ML system so far?

Interviewer 1

Okay.

Interviewee

So you said the quality issues in data model and

Interviewer 1

our system. Okay. Basically it's an open-ended question, so if you think of anything Yeah.

Interviewee

Uh, quality issues, um, I would say differ, uh, in industry differs from, uh, the classic, uh, machine, uh, machine learning community benchmarks. So, so what we've seen in industry is that, um, the, the images have, uh, Well, first I would say I, I would push in like, uh, you probably have a question for quantity, but I would push in quality in there also because they're, these two are really, really tightly bound.

Interviewee

Um, I would say, I mean, you can have a lot of data that is not useful at all. And uh, so I mean, whether it's for computer vision or Tableau data can have a lot of data. Uh, that can be, um, Actually unusable. So we've seen large database, uh, where like we're told by clients, oh yes, we have so much data. But, but the reality is that when you start, uh, looking into it, uh, uh, there's just not enough.

Data-low-quantity

Interviewee

And it, the more you, you have those restrictions, the more you ask questions and you try to build, um, the system, the less you have data left and then, uh, ends up, uh, um, with a very. Low quantity of, of data, but it's, at that point it's high quality for, for, for that machine learning, uh, application. So, so I would say, uh, it's hard to separate quality from quantity because if you have very these days with, uh, some models, if you have, uh, not many, uh, example, I mean, When I say not many, there's kind of a, a, a minimum threshold of data that you need, but if it's very high quality, you can probably have enough to, to get to a mvp.

Data-low-quantity

Interviewee

And then this puts you in a good position for, for scaling your project and getting the next, uh, uh, phase within a company. So, yeah.

Interviewer 1

Okay. So if I understand you correctly, what you're saying is there is a lot of. , but often time the quality is just not there. Yeah. So you end up with, when, when do you, you filter out data that is dirty or not usable, you end up with really not a lot of data.

Data-noisy-features

Interviewer 1

Yeah,

Interviewee

exactly. Okay. And, uh, yeah, and, and it's in computer vision. It's, um, I mean, data is not easy to get in, uh, for industrial applications. There's so basically, uh, I would say, I never say, oh, yes, okay. I never take for um, companies, At face value and they say, yes, we have a lot of data. Uh, to me there's like 10 questions that needs to, for me to determine if they actually do have the correct data.

Interviewee

So, so to I have the correct quality data and what will be left, I mean, to me, like it's, it's normal that after data is not usable, but, uh, we need to know what f is. If it's, uh, it reached, uh, the threshold to actually create a solution at the level that's needed.

Interviewer 1

I'm really, it's really interesting. And, uh, what are the 10 questions that you ask if you're able to tell me?

Interviewee

Oh yeah. So, um, uh, well, do you have access to the data and how long will it take to actually get the data? And, and, and that's usually, that's probably the, the most important and. What you, and, and to test this, you kinda need to ask, um, okay, can I, we have a sample of the data and, uh, sometime it's, uh, even that is really difficult because there's different ownership of the data.

Data-accessibility-issues

Interviewee

Also, there's the question, can it be shared outside the organization and so on. So though these are, I mean, I can go on, but like, yeah, it's a kind of, uh, uh, I kind of, I didn't write them down because I. Ask them every time. And like, uh, it's more like a fluid, um, discussion, but it goes around this like EC data, accessibility, um, getting a sample.

Data-accessibility-issues

Interviewee

Uh, and then, uh, yeah. Until you do that, it's really hard to, to really move on a project.

Interviewer 1

Interesting. And can, can you follow up on data accessibility? So you're saying sometimes they have data, but it's difficult to get it or use it for ML system, uh, for

Interviewee

project. Yeah. I'd say there's, um, within the company there's many people and there's many departments and, uh, so the.

Interviewee

The, say the innovation team or, uh, the people, uh, want to develop that model in that, uh, department doesn't have necessarily have access to, they know the data exists, but it doesn't mean that they can access it easily. And they make assumption about what can be accessed and, uh, who needs to sign off on giving access to that.

Data-accessibility-issues

Interviewee

So, so it's not, uh, it's not, um, even companies that say, oh, we're, we're fully digital and so on. I mean, they're, there's still. . Yeah. There's a difference between like, uh, Um, a big, a big server full of data that's disorganized and uh, that some people don't have access. And like, actually very well curated data set as compared to say in academia because each data set in academia has been really, well, I mean, you, when you, you write a new data, when you build a new data set, you basically, uh, compan it with the publication because cuz it's, that's the.

Data-accessibility-issues

Interviewee

Especially when they make it of open source, at least for the research community. Um, I mean, it's, it's a lot of work and it's a lot of effort. So, so I would say, uh, there's a reason why the academic data set, I mean, they're, they're good. I mean, they, they're good to prove some of their points, but, uh, but they're not industrial data set.

Data-accessibility-issues

Interviewee

Yeah.

Interviewer 1

I see, I see. And, uh, that's. , but for my interest, uh, why can, why is like data accessibility a, a problem? Why do they want to prevent some data to be, um, used by you, for example?

Interviewee

Oh, it's not that they want to prevent is that, uh, each, each data set is bound by certain, uh, um, okay. Where was it created? Was it created by, by the company you're, you're working with or by another company and then, you know, some, some companies.

Interviewee

They need to use the data of their clients. And that becomes a really pro problematic. So, so there's some kind of, uh, internal like, uh, data security issues basically, and also ownership depending of, uh, what it is. And, uh, yeah. And that's just, that's for, uh, accessibility. Yeah.

Interviewer 1

Okay. Super interest. Thank you.

Interviewer 1

Uh, so I will ask you a couple of questions regarding data collection and, uh, like I said, if you have experience on it, Please tell me about it then, if not, no problem. Uh, so the fir uh, did you ever use, uh, people that manually collected data for you? Uh, like Oh, for

Interviewee

annotation? Annotation? Yeah. Yeah. Um, no, we don't I mean, unless you, you, you we're talking about our clients that basically they, I didn't use them.

Interviewee

they gathered the data for us. Um, there's usually you need someone that will do that, but if you're talking about someone exterior, um, I know one company that did it, uh, but they're not an AI company. So, but, uh, yeah. Okay. So prob probably the question that follows. Did you ever use an annotation service?

Interviewer 1

Oh, no. Yes, but not right now. My next question was, uh, so, so your client, uh, what were the quality issues you had with the data? When the, uh, employee data collectors or labelers?

Interviewee

Oh, um, I think it's just, uh, I don't think there was issues, but there's definitely the, in that case, the annotation, um, Was not perfect.

Data-noisy-labels		Data-unnormalized-annotations

Interviewee

So, so this, uh, we saw that later on, you know, when, and we don't review like every piece of data, but when you really start drilling down, there's some like inconsistencies and, uh, and these are huge effect, basically the contradict the model, predict, predict the model training. So we really need to, to make sure the annotations are really correct and, uh, so yeah, so, so I would say if.

Data-noisy-labels		Data-unnormalized-annotations

Interviewee

If you do data annotation coming from the exterior, you just need to really have a way to, to maybe it's a very good annotation company, uh, so that, because errors in annotation are pretty, they, they will be what blocks you from going further and, and. And it's very expensive to, to go figure out what are the annotation issues.

Data-noisy-labels		Data-unnormalized-annotations

Interviewee

So

Data-noisy-labels		Data-unnormalized-annotations

Interviewer 1

Yeah. And what do they do to fix these annotation issues usually?

Interviewee

Well, that's the, that's the thing. Uh, de depends on which, where you are in the process, but I would say, um, you have to go back, . You have to go back and, and, and they need to be verified. So, so yeah. So I mean, Yeah. And that, that's again, like, um, that's why the academic dataset are pretty interesting.

Data-collection-external   Data-collection-public-datasets

Interviewee

And I mean, just re I think, I don't, I don't remember, but remember there, there was some correction lately in, even in ImageNet, at least for, uh, so there's, there's still updates in tho in those very, very widely use data set. So, um, so I mean, that's. It's important to take note of that. So imagine industry when it's created by x say two to three people as compared to a thousand people that looked at the data.

Data-collection-external   Data-collection-public-datasets

Interviewee

So, so it's a really different, uh, level of quality. So we need to be very careful. Yeah, that's for

Data-collection-external   Data-collection-public-datasets

Interviewer 1

sure. And probably the last question regarding this, this aspect. Um, how do your client detect that there is a problem, annotation problem in the. Uh,

Data-collection-external   Data-collection-public-datasets

Interviewee

well on earth side, we detect it from the, um,

Interviewee

well, when the, when, when you realize that there's, um, kind of a ceiling , there's some contradiction and there's, uh, certain techniques don't work for, um, to, to get this, the accuracy or the metrics up. And then you can see like, um, I would say, uh, Dissonance, like, uh, basically competing behaviors on certain, uh, certain items.

Interviewee

And, uh, so, so these you can see by evaluation otherwise. And then after this, the way, it's basically going back to the annotation and really correcting them and it's up to the, the data, the person that, as the ownership of that part of the pipeline.

Interviewer 1

Okay. So, so basically, After we built a model, we see that the performance is not what we expected, and at some point we realized that annotation, uh, was a problem

Interviewee

somehow.

Interviewee

Yeah, there's, there's certain, like, um, I would say, I mean there's certain set of hypothesis that you can conclude from, um, when models stop improving. and, uh, you always need to revert back to, uh, certain parts of the data to figure out like, yeah, what's really causing these issues. Mm-hmm. , because the architectures are pretty solid, well, I wouldn't say pretty solid, but they, you kind of know what to expect from, uh, a model architecture and, uh, when you don't get it, uh, then the really, and you have to be careful.

Interviewee

And also when you get something too good, you also have to be careful. So, so this, uh, . I mean, uh, yeah, that's a, that's a direction. Yeah. Thank you. Yeah. Be because I mean, you always have to protect against overfitting, you know, so, because if you, you get a hundred percent on your training set in mean, that's, that's the worst thing you can ever get.

Interviewee

I mean, you need to care about the validation and even then, You need to be even careful with, uh, validation. I would say, uh, that even the way that I, so, so that the way the data is handled is really important. Uh, and so that you don't mix up things cuz then you, you get like, don't, you should never.

Interviewee

Training data with, in the evaluation with the validation data, and, uh, it's, it's easily, it can be easily mixed up on the data sign. Even the best data scientists in the world will need to really be careful, uh, because, you know, uh, when use the code, basically you do a split and then, uh, if you save a checkpoint and you, and you, you retrain from there and then you shuffle.

Interviewee

Or it shuffled because you didn't put the right flag, then that means you're mixing up the training data with the validation data. So I mean, there's, there's a lot of, uh, so what I'm saying is like, data needs to be read carefully. Data adaptation.

Interviewer 1

That's, that's such a good point. Data leak always an issue.

Interviewee

Oh, uh, the, you wouldn't believe the, the, the danger of data leak. Uh, you. You need to understand really, um, the client data when you receive it, you need to really understand it very well, uh, because to prevent data leak, . So, and, and not even that, but like, it's not even from the data science point of view, but just what's in there, what's in the, what's in the, the tables.

Interviewee

You need to, to know exactly what's, how things were given. And that's not always, uh, accessible. So you need to really have those discussion.

Interviewer 1

Yeah, and that's, that brings another point I find it interesting. So you said, uh, oftentimes you don't have the documentation to understand your data. Yeah,

Interviewee

exactly.

Interviewee

Yeah. Mm-hmm. Okay. So that's, that's why that then I have my eight next, next questions related to everything in the data. So yeah. ,

Interviewer 1

thanks. Alright. Uh, we'll move on to, uh, data preparation. So we talked a lot about collecting data and uh, yeah. Any issue that follow us up from that? Uh, yeah, let's talk about data preparation.

Interviewer 1

Yep. Um, which data type have you worked?

Interviewee

Um, time series stabler and, uh, imagery videos, like all of them, I guess. .

Interviewer 1

Mm-hmm. And have you encountered any issue with these data type at some point? What do you mean by issues? Uh, quality

Interviewee

issues. Yeah. For sure. All them. , I mean. Yeah.

Interviewer 1

Yeah. Perfect. Thanks. Um. , have you ever measured the quality of your data and or tried to improve it?

Interviewer 1

I guess yes, but yeah.

Interviewee

Yeah. We, we, we do that all the time. Uh mm-hmm. , there's different ways. I mean, uh, uh, but I would say like the, one of the key part is really getting, analyzing, like as a human, your data understanding, like, uh, I know sometime we don't have the documentation, but making sure we have access to that and that gives you really an insight to really what's happening.

Interviewee

Um, we wanna skip that most of the time and automate all this, but like for, for a project to be successful, you really need to basically understand the data. And I mean, there's the concept of data scientists where we think it's, um, , the job of a data scientist is take some data, just move things around, push a few aug, push it through a few algorithms.

Interviewee

But I think, um, That's, maybe that's the job of a data scientist, but a data engineer would need to really understand, like, I, I don't know. The term doesn't really exist and it's like a, um, but really to the data engineer is like someone I, I think a position that doesn't really exist, but it's someone that would be in between the data scientist and the domain expert, where you basically, uh, get this, uh, this documentation going and we understand like, uh, uh, what polls what.

Interviewee

Yeah. So, so that's, that's, I think, uh, that's probably the most important. But it becomes a very, uh, uh, it's, you cannot automate this, you know, so, which is, uh, so it's really getting, uh, that knowledge. Mm-hmm. .

Interviewer 1

So my next question was do you have tools to help you clean and, uh, you're data, but from what I underst.

Interviewer 1

More. Yeah,

Interviewee

yeah, yeah, for sure. We have plenty of tools for that. Okay. And, uh, we, um, yeah, and there's a, a lot of open source, uh, I think for tabular data. So many things to help you clean the data. Um, lots of open source, uh, to help you do that. And, um, and image imagery. It's, uh, say, um, it's not, uh, it's not that bad.

Interviewee

It's not. imagery, kinda take them as they are images, so, so that's not necessarily needed to clean images.

Um, it's, it's part of different, uh, data pre-processing that we use on the, depending on the project and the

they are. Yeah, exactly.

Interviewer 1

Okay. And can you name a few tools you use?

Interviewee

Um, well, I don't use. All myself, but we'll use, uh, different, um, uh, tabular data evaluation just to know like if

they're missing values.

Interviewee

I mean, uh, you know, it's just panda manipulations. Uh, and for, for the tabular data. And, um, and for

imagery, I'd say it's like, uh, self written codes. If you just, uh, split the data correctly and like, uh, to, uh, I

mean in that case it's, uh, pre-processing, PyTorch pre-processing, and uh, well, you know, the data loaders

and so on.

Interviewee

So that's, uh, those are, I mean, the, they're already tightly integrated with, uh, with KRA and PyTorch. So, so

that's, that's I would say so, yeah. I see.

Interviewer 1

Thank you. Uh, moving on to, uh, oh, well, is there any other data quality issue we missed that you consider

relevant data

Interviewee

quality issue?

Interviewee

Yeah. No, I don't, uh, yeah, not thinking of something at the moment. Okay. Thank you.

Interviewer 1

Alright, moving on to model evaluation. Uh, how do you evaluate the quality of your models?

Interviewee

Uh, well, there's two ways. There's the. Uh, the raw way, like, uh, that's the data scientists like, uh, numb, like evaluations from the, from the pipeline, but we also have a tool for doing that. So we have, uh, one of our products is exactly is doing for model evaluation. So, so that's what we use internally. So I mean, we built our own tool and we sell our own tool just for that.

Interviewee

So,

Interviewer 1

yeah. That's super interesting. Would you like, like to talk about.

Interviewee

Uh, well, I mean, I can talk about what's on the web, whatever we make available.

Interviewer 1

Yeah, for sure. Yeah, sure.

Interviewee

So do you have question related to that to guide me or, uh,

Interviewer 1

so how do you evaluate the quality of your model with your tool in general? So,

Interviewee

so basically we test the robustness of the model and, uh, we pass it through different series of tests to figure out, uh, , what are the weaknesses and, and also what are the, uh, what's working?

Interviewee

And, uh, the idea is like to, uh, to not on, to, to get lots of good information that is not just like the validation, uh, Accuracy or whatever other metric it is for that model. So, so that's basically what we've been doing. And if you want to, you can go on the website and then you can see a little bit more.

Interviewee
And, uh, yeah. And if you have question, then can

Interviewer 1
definitely see it. Yes. Yes. I, I, I seen it. I, I listen to video people, it's just for, for the interview, so. Okay.

Interviewee
For sure. All right.

Interviewer 1
Um, have you ever used existing benchmark models or quality aspects to evaluate your model? In which way to, uh, so if you use a model to evaluate the quality of your model, uh, uh, you compare your model against another one to evaluate its quality,

Interviewer 1
uh, a pre-trained model, uh, any.

Interviewer 1
Uh,

Interviewee
uh, well, I'm not so sure to answer, but, uh, we never just build one model. We build, we build like, uh, a Plato of models. We, I build like, um, , uh, we always pit those results against each other. I mean, we just, we, we do, we won't just build like, uh, I dunno, I'm saying, uh, so, so we'll, we'll train different type of, we'll train.

Interviewee

Okay. So assuming we went through all your first questions, like we have a, a good data set, I said fair amount, or even if it's not a fair amount, that's what we have. So we have to work with it. So what we'll do is we'll run, uh, different models on the, we'll train different models on that data set, and then we'll keep the one that have the best, uh, characteristic at that stage.

Interviewee

So, so, yes. So in this case, but, but if you say comparing with the benchmark, uh, I would say we, we know from different papers, what are the state, what's the state of the art on certain data sets. So we can evaluate like, um, but that's more like, Because it's not the same data set, so, and it's not the same characteristic.

Interviewee

So, so it's hard to evaluate like an industrial project with a smaller data set compared to something that trained on the sci far, you know, so , I mean, which is, uh, very low resolution. And, uh, and it's just almost for, it's, it's just benchmarking at that point. Benchmarking the power of model. So, but yeah, so.

Interviewee

So I don't, I don't know if that's, it's, it's not really worth taking an ImageNet model comparing, uh, a pre-trained model because the data is never going to be ImageNet, you know? So otherwise, we'll, we'll just kind of, that means we're just using ImageNet, uh, an ImageNet trained model, and I don't know why , we're developing this really complex solution.

Interviewee

So Yeah,

Interviewer 1

I agree with you. It, it's a good answer. Yeah. Alright. Um,

Interviewee

just, I, I have a question related. So, so when, in addition to that, when, when evaluating the model, some people uses another model to evaluate a model. Like the model would be an, for example, if, for example, you wanna evaluate your business, you can train a model to evaluate their business of another model.

Interviewee

So some people started to using those techniques. Uh, do you use it in your, uh, in your, uh, pipelines? Um,

Interviewee

Hmm. Let's see. I am aware of, of that and I, uh, we did experimentation. I mean, it's, uh, And I think I can say more at this stage, , but, but yes, we're, we're aware of, um, the use of models for, for determining, uh, uh, for evaluating models. So definitely, um, I would say that,

Interviewee

hmm, I'd say it's, uh, sometime very overkill. So, It's, uh, but it's, it's, it's, uh, it can be good techniques. Yeah. Okay. Thank you. I mean, , to be honest, like, say the data , what people really do, they just look at one number, which is, uh, like, uh, you know, if you're, if you're lucky, people will look at the F1 score and, uh, accuracy and so on.

Interviewee

And, and that's, that's a long way to go to start explaining that. Oh, yes. says, uh, says, uh, says this extra things. It's, it's every big technology needs to be, um, uh, when you use a pretty intense technology, you need to make sure that the user will understand also and then mm-hmm , and that can be really, uh, challenging.

Interviewee

So, so we just need to have. The right approach. Yeah. I see. So, so generally the clients are mainly interested in the accuracy and those clear metrics. They, they're not really interested in the ones that are Oh, no. They, they don't, they need to be taught what a metric is. . Okay. They need to, that's why I'm trying to, so, so they need to, to be educated on the, how, to understand the model correctly.

ME-metrics-meaningless-for-stakeholder

Interviewee

Mm-hmm. and, um, and sometime they, you know, when you come in new projects and they, they've been using the wrong metrics, so you need, you need to bring the clients to the new metrics or the, the actual, the metrics that matters. So, so, so, yeah. So it's like, I mean, maybe I'm going back to, to the earlier is that there's a really, um, Each project needs to be, uh, looked at closely.

ME-metrics-meaningless-for-stakeholder

Interviewee

You know, it's a, it's a, it's like there's not, um, for AI services, there's not like a, you just can just run it through the system. Everything's gonna be good and everyone's gonna understand. There's a, you need a for, for projects like this, especially when you're joining Project, to help them, uh, get unlocked basically the, what they have and, uh, We really need to work on the metric understanding and Yeah.

Interviewee

Mm-hmm. . I understand.

Interviewer 1

Thank you. Um, have you ever accessed the quality of ML model prediction with users of your system?

Interviewee

Users of my system? Uh

Interviewee

uh. Okay. Can you make more context? I, yeah.

Interviewer 1

So. For example, imagine you have a ML system and this ML system as user. Uh, did you ever, oh, sorry. Did you ever try to assess the quality of the ML system with how your user perceive, uh, your ML system? So, user acceptance test, this is what I'm asking.

Interviewee

Oh, but that's, um, I mean, that's a ui ux question in a way.

Interviewee

So it's like, Is it getting feedback from users? Is that what you mean? Yeah.

Interviewer 1

Yeah. But well, it's more about how they enjoy the, um, I will give you an example. So, uh, let's say a

consulting firm in ai mm-hmm. , uh, they can build a AI model and just consider accuracy or, or F1 score. But

oftentimes something they do is they put the product in the hand of the.

Interviewer 1

The client and they see if it's like what they expected, and this way they, they discover quality issues. Uh, so

what I was asking is if you ever did that and discovered quality issues.

Interviewer 1

Oh,

Interviewee

I mean, that's pretty much what we do. So I mean, that's, uh, that's the, the whole point. And if we're, uh, ,

there's different ways to put things in the, the hands of, uh, the users or the clients. And, uh, I'd say that's,

that's kind of, I mean, it can go from, uh, presentations. Okay. You know, that's, that's where, when there's

meetings, that's really when we are trying to get really.

Interviewee

Really good feedback, uh, because then you have the, their full attention and you can also ask those

questions. But also for them to interact with the data is, is, is also really important. Uh, just in.

Interviewee

Uh, okay. So, yeah, so that's, I guess, uh, I'll finish my sentence, . Okay. Very

good.

Interviewer 1

Thank you. And what are the quality issues you encountered at that point? Usually quality

Interviewee

shoes. Okay. In, uh, like I said, annotations, uh, you can have like, uh, Different type of formatting. Uh, the wrong data set that was used or sent to us, uh, can be, um, uh, data leaks.

Interviewee

I mean, that's, that's also, uh, so that's not always easy to identify, especially when it's from the u uh, the client side. Um, it's, there's a, so I would say. Also, um, uh,

Interviewee

there's a kind of, uh, inspector kind of work to do there. So esp uh, with, with clients especially. So you basically, it's a, it's like a, like I said, I, a series of questions and you're trying to really, uh, figure out like, uh, what they view, what they understand, and, uh, you're trying. Because you'll learn things that you never expected.

Interviewee

So it's that interaction that's important in the, yeah. Okay. Perfect.

Interviewer 1

Thank you. Uh, is there any other issue with model evaluation, uh, that you'd like to talk about?

Interviewee

Um, yeah, I think, uh, yeah, no, I, I think I'm, maybe we can come back, if I think of something.

Interviewer 1

Um, I will skip model. I have some questions about model deployment, but I will skip them since we are running out of time. Okay. Uh, I'll jump to model maintenance. Uh, so how do you, and if you have expertise or if you don't have expertise in this, uh, feel free to tell me.

Interviewer 1

We're gonna move on to the next section. Mm-hmm. . . So how do you ensure that the quality of machine learning software system does not decrease over time? Oh, you do a good job first,

Interviewee

and, uh, but, but it is, it's, it's kind of a joke, but it's not a joke. . So, so if you did a, a nice, robust model, then you'll get, uh, something that works better for.

Interviewee

and, uh, and then it's all about monitoring and, uh, and that's all trade in itself. Um, there's a, what we know is that there's a lot of monitoring companies out there, and they all have their strength and weaknesses. Uh, there's monitoring software, uh, there's open source monitoring solutions and so on. But, uh, that's a, that's a huge topic, I would say.

Interviewee

But, uh, but that's, um, It's a big space, you know, , so,

Interviewer 1

yeah. Could, could you name a few, uh, tools you maybe have used? Uh,

Interviewee

uh,

Interviewee

I mean, clearly, um, Onix is a good one. Uh, but that's not for monitoring, but that's like for, uh, production. Uh, ML flu is a good, good tool also. Um, but I would say,

Interviewee

It's, um, so at this stage, uh, no, people claim a lot of things, but there's not that many models in production. So, so that's, uh, uh, banks do have a lot of models in production, but they have their own, uh, internal, uh, serv services. So, so to me, uh, We're not quite there yet. , but I mean, it's, uh, it's moving. So I mean, um, I, were not in each of these project need monitoring, but uh, as we progress in the AI space, I would say it's, uh, it's not quite there.

Interviewee

So, so yeah, so we don't use a lot of these tools and uh, I mean, we use our tools internally for, in production for sure. Um, and we have our all internal system. We have, uh, a w s, uh, create reports and so on. So we have all these these, but, uh, for machine learning, I think, uh, uh, we don't work too much in that space at the moment, so, so I guess maybe whatever I'm saying is not so good.

Interviewer 1
Perfect. Thank you. Uh, I see, uh, usually we have the, uh, premium version of Zoom. Uh, but today one of our team member couldn't be there. Mm-hmm. . Uh, so if you, if you don't mind, we have maybe like 10 more minutes for the interview. Sounds good. No problem. Okay. And I will send you the other link right now before it close.

Interviewer 1
Oh, okay.

Interviewee
Let's.

Interviewee
Is she ? Okay. Yeah. I'll, I'll see you in two minutes. Perfect.

Interviewer 1
See you. Thanks.

Interviewer 1
All right. Thank you Interviewee. All right. Uh, so just to follow up with your, the last point you said, I think it was interesting you mentioned that bank often have a lot of system ml, well ML component that are deployed. , uh, why is that, that bank have a lot more than compared to other companies? Oh,

Interviewee
okay. Uh, well that's a really good question.

Interviewee

Um, they not necessarily going like it, but , um, they, they did that, uh, a long time ago, , so, and the, but they're, they're like older models and, uh, so they do have things in production, but, um, from what I've heard is, uh, Um, it's still hard to get there. I mean, you can update those models and so on, but they, it's, it's not easy to get actually there.

Interviewee

Uh, so for people, so I mean, there's, I mean, we're not necessarily talking about like huge enterprise, but, well, yeah, like, no, what I'm talking about, there's, there's like huge user organization. So I, uh, you know, whether it's, uh, I can say names, but, uh, these organizations have some things in production, but it doesn't mean that, um, lots of things are in handle internally.

Interviewee

And, uh, so that's for one thing. And for the other thing is that, uh, you, you don't necessarily have to assume that it's the most modern stuff that is being used in production and. So they even, uh, adjusting the most re to the recent stuff is, is still a challenge inside. So, yeah, so, so to summarize, uh, the statistics that you see, like 85% of projects won't end up in production from Gartner.

Interviewee

Uh, I don't know, only, uh, 10% of, I dunno, 15% of companies actually. I have models running in production. Uh, it's the, these are all true, this, that's the reality of ai. So, so we're in this, um, uh, we're in the reality check phase. So there's the hype and then, you know, from the, from the cycle, we're in the reality check phase.

Interviewee

So, um, and there's, uh, hurdles there. So, so that's, uh, that's why I usually go back to, okay, you need to really.

Interviewee

Look at the data, understand what's happening, and make sure everyone's understand. Okay. And then, and, and sometime there'll be things you'll learn and then you'll have to adapt it. Yeah.

Interviewer 1

Okay. And why does these project fail? Is it because they don't have the expertise? Like is it a logistic or knowledge problem or It's more, AI could not solve

Interviewee
this problem.

Interviewee
Uh, it's a logistic, uh, problem. And it's politics and it's so on. So I don't think it's about AI in some problems. I mean, it's, uh, I'd say there's some, I mean, if you have good data that AI should be able to do, uh, really good, but. You don't have good data. So , so then, um, I mean, it's not the models, architecture is necessarily the problem.

Interviewee
It's, um, I mean, it is sometime because if you wanna reach a certain level, right? You need like, uh, the very, very cutting edge, uh, architectures. But, um, they're really powerful. The models, you know, like I, I mean, , I mean, if you can handle like 1000 classes with, uh, and get all the right bonding box, even though you're not totally perfect, and it's pretty amazing.

Interviewee
You know, like if you look at, uh, things like yolo, uh, , I don't know. YOLO 15, I'm joking because there's no YOLO 15, but they will be , I think we're, they're like, uh, there's V five and now there's maybe V seven, and I don't really follow the terms because some of them are just like a joke. But, uh, yeah, . But yeah, what I'm saying is that they're, they're, they're good.

Interviewee
Those models. I mean, they, they do a really good job and, uh, if you have the right data and the right annotations,

Interviewer 1
Yeah, the, the early papers of you were, uh, pretty funny. That's . That's just to say. Yeah. Yeah. All right. Thanks. Uh, yeah. Have you ever added the issue, so we were talking about model maintenance.

Interviewer 1

Uh, do you have any quality issues, other quality issue, uh, you didn't mention but you'd like to share with us?
Uh, in model maintenance? , yeah.

Interviewee

Monitoring. Yeah. Like, like I'm saying, we're not really fully working on monitoring, so, so I mean, uh, I mean, I, I can know about the monitoring market, uh, very well, but I would say, uh,

Interviewee

I, um, it's just you need a nice. , uh, you need to be able to get your, to iterate, and that's what we've seen. It's important to be able to iterate, uh, at the right pace. And that's not necessarily, uh, but I even to get, to get there to deployment is, is, is kind of, uh, that's the challenge where everyone's happy.

Interviewee

Otherwise, I mean, you can deploy the things that don't work and then you'll lose confidence. . Uh, and then that's, uh, pretty much the, the end of your project. So . I see. So, yeah.

Interviewer 1

Yeah. Uh, I'm, I'm thinking about it regarding something you said earlier. You said, uh, ML project, they failed because of, um, sorry, uh, logistic and, um, I, I forgot this, this again.

Interviewer 1

Um, about data politic,

Interviewee

politic. Politics. Politics, yeah. Yeah. Politic. Yeah. I mean, , there's many people involved in the project, you know, so, so that's, uh, can be like, oh, uh, there's, I mean, there's someone new in the project and then, , the project doesn't continue. So there, there's a lot of things involved in the ML projects.

Interviewee

Okay. Well, politics, it's not like people fighting, it's just that there's a change of, uh, people are, I mean, there's movement in the AI space, so people get offered jobs, then they go somewhere else, then the project needs to be taken over and then things change, things slow down. And, uh, so, uh, yeah, so, so that's, uh, those are, are things that really makes, uh, those projects.

Interviewer 1
Yeah. Would it be fair to say that, uh, they have like a AI project? Well, no, maybe not. I will say, I was about to say that they have a low return on investment, so they just, just drop the project. But it's more than just return on, on, on investment. It's, uh,

Interviewee
yeah, it's more, it's more, I mean, the return on investment, I mean, it needs to be demonstrated for, for each of those projects.

Interviewee
For these. to be even, to continue for sure. Um, but also it's just like, uh, priorities change and uh, and sometime like, uh, I mean I've seen places like, uh, someone seems to know better , so, so, so, so it's there, there are some politics, some, some groups wants to take over, uh, the project because they wanna dabble in ai and, uh, and that's, uh, You know, so, so we've seen that.

Interviewee
I mean, it's, uh, when, uh, it's not often a very good, uh, approach, uh, because like I'm saying, AI is, and data science is, uh, is so much more complex than we think. You know, it's, it's really not just like taking like, uh, like, uh, like, uh, running your tutorials and, and running a model. You can't do that, but you really need, uh, the expertise, uh, to really know, understand where the project's going, why it's going like that, what to do next, and, uh, how to unlock all these steps.

Interviewee
So, so there's, uh, I mean, my point of view, like am I manage over 15 projects and, uh, that doesn't even count the projects we did internally. So, so those, uh, so I, I've seen so many types of. Things. And uh, and it's just a, a project needs to just be, uh, you need to develop a very nice relationship with the clients and, and

foster that.

Interviewee

Mm-hmm. . I see.

Interviewer 1

Thank you. Uh, probably some of my last question. Uh, what are the main. Robustness or explainability, uh,

issue that your client have generally, uh,

Interviewee

well, what models are not robust? And that's pretty much the biggest . That's why we build those tools.

Interviewee

So, so, uh,

Interviewer 1

yeah. Do you think you can be able to give an example or not really?

Interviewee

It's super, I'd rather not. It's a little bit, uh, it's, there's too many, uh, NDAs in place to, to give like in precise

example. Uh, yeah. Just I wanna be careful on that part. Yeah. Thanks. Thank you. Thanks for . Thanks for

giving me a way out.

Interviewer 1

Yeah, no worries. Um, did you ever, ever encountered fairness or scalability or privacy issue? Any other

quality issue?

Interviewee

Uh, privacy for sure. Uh, uh,

Interviewer 1

fairness. Yeah. Uh, so fairness, is that a. Sorry, what was that? A

Interviewee

question? No, no. Um, you're saying, so you're, what do you mean by fairness? I mean, your definition. Uh,

Interviewer 1

so a model that does not discriminate certain group of people.

Interviewer 1

So for example, women or people of, uh, skin, different skin color, things like

Interviewee

that. Like that, that, yeah. Models will discriminate. Uh, they, they will be, unfair . And that's, that's their, it's

that, that's the data that says it. So for example, um, so if you have a say, um, uh, like a very well known like,

uh, open source model, uh, so for home credit, so you can see that model will always favor women over when

you, you're giving a loan to someone.

Interviewee

And, but that's, that's because. They will, there's more chances that they will repay the loan. So, so the model

does discriminate. I, I mean, it doesn't discriminate. It takes the data and they're in, if you're a bank or an

insurance company, that that's what, what it does. You know, it needs to figure out, like, I mean, , you pay

more insurance.

Interviewee

You, you pay more insurance on your car than, than the woman may Same age, same, same job, same. So,

so that's, that's, that's the, that's a, that's the real, the sta statistical reality. So yeah, that's, and I, I know, I

know the, they find terrible dis I mean, those, those really make the headlines, but it's the data that.

Interviewee

that created this, the model was kind of, it's not the model discriminating, it's a, it's, it's, it's what people have

been doing forever. , you know, so it's the people that are, that are discriminating because the, but you need

to be really aware of that. And in the data preparation, then you can, I mean, it's, it's re it becomes a choice.

Interviewee

I mean, uh, so in, in the case, on the home credit, you would like, I mean the model needs to do. Goal of the company is to, um, make it fair than for between women and men. Then, uh, it's a really different way to handle, uh, the data. So basically you, you would make sure that, uh, um, That, that value, whether male, female, would not be taken, uh, male, female, or other, would not be taken into account.

Interviewee

And, and that's, uh, but I don't, that's, uh, like, oh, that's totally a political decision within an insurance company, which is, uh, so, um, so yeah, so, so, so that's, uh, , I think it's just like the data's there. If you train on historical data, you will replicate the same biases. Exactly the same thing. I mean, a model is not magic.

Interviewee

It's uh, it's basically, uh, it repeats the history that you gave it. So, so if you want to, uh, avoid racism, then you need to. It and do the same historical data because you don't have other data, whether you create new data or use that data, but you handle it really carefully and, uh, and the right way with, so you can make your models say what you need to say and, uh, and how you want it to say so, so it becomes like a really a decision.

Interviewee

So, yeah, I mean, because it's, it's not gonna do anything new, you know. We think it's make new prediction, but it makes the same prediction based on the similar data. So, so it'll replicate our, our terrible history .

Interviewer 1

Yeah, yeah, yeah. Sure. That's a good point.

Interviewee

Thank you.

Interviewer 1

And, um, my last question, uh, in your opinion, what is the most pressing quality issue researchers should try to solve?

Interviewer 1

Quality issues?

Interviewee

Uh,

Interviewee

like the most pressing. Like from what? Like, uh, you have a, just trying to, I mean, there's,

Interviewer 1

yeah, there is no good answer. It's just, um, What is a, a pain point that, uh, maybe you have in your daily life and, uh, it's totally great to not have it

Interviewee

anymore. Whatever foot's working on , you know, to be honest, like, uh, that's, uh, I think, uh, those are the things he knows and, uh, and, uh, and I, I agree with what he's doing.

Interviewee

So , so I would say yeah. It's not, it's not to be nice, but, uh, cuz I know him and, and it's just what he's working on is really valuable. So, I mean, what they did at Simo was, uh, this is the kind of things that are needed for, like, I, like I said, like to avoid the next, the next AI winter. I don't think it's gonna be a winter to be honest, but it might be like, uh, like, uh, Climate change, winter, you know, so , so it's gonna be a warm winter or, or warm or like a fall, you know, like, uh, so, so, so I, I do think, uh, yeah, I think it needs to be dealt with,

Interviewer 1

you know?

Interviewer 1

Uh, cool. Down just a bit

Interviewee

more. Yeah. Cool. Okay. That, I like that. So, so we want to avoid the cool down. Um, we are in the period of cool down. Uh, this, this is just a reality check, uh, but. I mean, there's not many groups that like you guys that are actually doing this. So this is, this is what's going to prevent that know, so, so if you continue on doing what you do and that's the going to actually help, uh, that become a cool down.

Interviewee

Yeah.

Interviewer 1

Perfect. And before we close that you have any other qualities, uh, you tell about and you forgot to mention, or is it. Or,

Interviewee

yeah. What do you mean by quality issues?

Interviewer 1

Uh, uh, so, uh, anything, yeah, any quality issues with machine learning, software system. Anything maybe you didn't mention yet?

Interviewee

Uh,

Interviewee

yeah, I mean, I, I mean, I talked about robustness, uh, data quality. Uh, I never said model quality, so, um, I'm pretty, uh, I'm pretty happy about the architectures of models, to be honest. Uh, and every, well, I mean, yeah. Yeah.

Interviewer 1

So I almost a solved problem model

Interviewee

architecture. Well, I mean, it's solve, it's not solved, but it's,

Interviewee

See the other problems are so much bigger, so , so, so at least for, for industry. And, and I think, yeah, I mean, to be honest, what, what your group is doing is, uh, it's really is on the right path. So, yeah. Perfect.

Interviewer 1

Well thanks a lot Interviewee to, uh, taking a bit of time to share your expertise with us. I think it'll be, uh, really good our or our study.

Interviewer 1

Uh, so. I agree. It's, uh,

Interviewee

it was very good, uh, interview. Yeah. So sorry, I, I I should be like all about the hype and everything, but, but we know the reality and it's Yeah. That's what, that's what we search. Yeah. Yeah. Exactly. And, and I mean, honestly, I find, yeah, when, when I learned about what you, what you do, and I was really happy to be similar because it's, it, it's.

Interviewee

That's what's, that's where it's happening. That's the, but not many people know. So that's . So there's a part where, but it's, it's, people are slowly waking up. I would say. So. So you it's uh, it's still slow. It's, it's slow. Yeah. Definitely. There's still a little bit of the hype left, but the reality kinda kicked in I'd say maybe, uh, the last eight months.

Interviewee

And there's a lot of talking. But not a lot of doing at Uh, so that'll probably last for next year and a half . So, which is, uh, yeah. So well keep working hard. And, uh, yeah. And keep me in touch if, uh, with your progress, uh, especially about that study. It's really interesting and, uh, perfect. Yeah. Really

Interviewer 1

well, we'll make sure to send you when we're finished, uh, and even before maybe.

Interviewer 1

Yeah.

Interviewee

Okay. Yeah, sure. Definitely. So, well, thanks so much. We happy to open. Thank you. Thank you. Well, all

right. Take care. Bye. Bye.

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