Interview 20 - PO

All

Interviewee

right.

Interviewer 1

Perfect. Uh, so to begin, could you give us a little bit, tell us a little bit about yourself. How many experience,

how many years of experience you have in, in machine learning or in general?

Interviewee

Mm-hmm. ? Uh, sure. So, um, I'm currently working as a lead data scientist for Company Y. Uh, I've been

there for about four.

Interviewee

Um, prior to that I worked for a startup, um, called Company X that was based outta San Francisco. Uh, and

that was like a cardiovascular health startup. I was there for three years, so I guess, uh, seven years total in

industry. Um, I also have a master's in neuroscience, um, where I did some, you know, I spent three years of,

uh, grad school kind of doing predictive modeling, but um, in a more academic setting.

Interviewee

So maybe 10 years total. Okay.

Interviewer 1

Perfect. Thank you. Um, so let's start with the first question. Mm-hmm. , what are the main quality issues you

have encountered with your data model or system so far?

Interviewee

Uh, yeah, a couple. So with data, um, I think there, you know, the types of issues I've encountered are maybe like impossible values, for example, like if you have the, you know, age of a, of a individual that's 200 years old or 300 years old, or negative 50 years old or something like that.

Interviewee

So, you know, sort of physiologically impossible values. Um, I think, uh, anything that's kind of dependent on. Like upstream sort of ingestion processes, like, um, we'll, we'll often see like delays, um, in, in ingestion processes that can throw off sort of downstream processes. Um, uh, what else? Um, duplication is another big one.

Interviewee

So like, you know, finding lots of duplicate values in our databases. Uh, I would say l latency and, and data lag can be one. Um, so not necessarily, you know, not having access to things that happened a month ago until, you know, a, a day ago or something like that. Um, so yeah, say impossible values, uh, sort of delays in like ingestion processes, duplicate values.

Interviewee

And then what was the last one I just said? Oh, just like lag between when, when an event is recorded and when it's available. Um, I would say are, are probably some of the common ones that I've encountered. Thank you. I have a follow up question. Sure. Since you mentioned the delay, do you have any problem, uh, for real time system?

Interviewee

I mean, you are, you are supposed to make a decision in real time, but those lags make it difficult for you to make the decision. Yeah, uh, exactly. So especially like, you know, we have the ability to, um, uh, you know, text people or email people, like instantaneously, but, um, the information that we're getting on those individuals might not be, you know, might not like work at the same scale.

Interviewee

Like even, you know, there could be several days or several weeks, you know, between when, when an event occurs. You know, maybe we would like to act on that event as soon as it occurs, but because of that data lag, we can't act on it immediately. Um, uh, another example is like, so like I, I'll refer like in healthcare data like this, this could be specifically like when a patient interacts with a healthcare provider, um, it might take.

Interviewee
You know, days or weeks or months for that healthcare provider to submit a claim or something to an insurer or, or otherwise create a record that, of that interaction that would be available to researchers. Um, I think this has a, a big implication when you're training models because, um, you might set an index date for your, like, features that you're observing.

Interviewee
Um, And that data will be a hundred percent complete because you've, you've already waited, you know, a certain amount of time for that data to become available, but then when you actually score that model in real time, you're actually working with only a subset of the available. Of the data that's currently available.

Interviewee
So your, the data in your, um, scoring window doesn't match the, the data in your training window because of that lag. Um, so you have to like introduce these artificial lags, uh, into your training process to try to mirror, um, what data might actually be observable. So that, that's kind of one example of like how it affects like a real time, a process that we're trying to implement in real time.

Interviewee
Does that answer your,

Interviewee
Surely you have already a answered the question was only follow up question. Thank you. Okay. . Cool.

Interviewer 1
Yeah. Uh, and may I have a follow up question on the follow up question? Sure. So how you introduce noise in your, well that noise was, uh, the lag in your training data set.

Interviewee

Yeah. I would say if we have a reasonable expectation of.

Interviewee

How long, you know, how long it would take for us to be able to observe, you know, 99% of our data or something. I mean, if that is a month, let's say, then when we train our models, we would, we would like stop. We, we would basically only use data. Um, we, we would basically ignore a month of data. So it's like if, if we're using 12 months of features to predict an outcome in the next three months.

Interviewee

That 12 months of features would actually be like minus 13 to minus one month. We would have like a gap month, and then the prediction would be the next three, you know, the next three months or something like that. Um, another way to do it is to like, uh, randomly like aate features, like based on how, on how you would expect the delay process to work.

Interviewee

But, um, I, I think that might be unnecessarily complicated. Um, I see, I see. Mortgage case.

Interviewer 1

Okay. And how do you detect the, uh, these, these issues? So the, the the like between what the like and your data, how are you able to see that you have a temporal data leak? Is it something you discover because you have a poor model idea or something?

Interviewer 1

Like how do, how can you detect it? Yeah.

Interviewee

Um, I think it's a known issue with this particular form of data ingestion. Um, I mean, I think it comes through observation, but. Uh, you know, it could be due to, yeah, it could just be delay, delays in, like, you know, again, if it's a, if it's a healthcare record, it could be a delay on the provider's end.

Interviewee

Um, or if we're dealing with like, you know, if we're ingesting wearable data, for example, like it might take, you know, one or two days for us to be able to observe. Wearable data or, or sometimes data is dependent. If, if we're using an application, the data transfer might not occur until the next time a user logs into their application.

Interviewee

So, you know, that might, that could be like several days to a week. Um, so I think these are kind of, we, we discover them by like sort of. Yeah, a lot. I don't know. It's like we, we don't necessarily have like good monitoring processes in place, so it kind of becomes just like a general sort of knowledge of how these systems work and, and they get, you know, documented and the solutions get documented as well.

Interviewee

Okay.

Interviewer 1

I see. Thank you. Um, did you ever, yes. I will ask you a few question about data collection. Yep. Uh, did you ever use the service? , did you ever use a, a medical data collector to, to get training data for your dataset?

Interviewee

Uh, at my startup I did. Yeah. We were running sort of a. Uh, something of a, so yeah, we , we had like a physical hardware that we used, uh, in a, in an experiment context.

Interviewee

Um, so we would have people come into our lab, um, and by lab, I mean a garage in Silicon Valley, um, and uh, for perform like a physiological experiment with the hardware that we had, um, available. Um, so that was, yeah, definitely an example of a manual data collection. . Okay. And

Interviewer 1

what were some of the quality issues you encountered?

Interviewee

Um, right, so it, I would, in this case, you know, we're measuring a physiological signal. Um, so in this case, like there, there are certain, certain factors that would affect, it would be like the, you know, strength of the connection between. You know, the human body and, and the instruments that we're using to collect.

Interviewee

So like, you know, people who have, you know, a lot of body hair or, uh, actually I worked in e in a EEG lab when I was, um, uh, in grad school. So we, we would, you know, put electrodes on the scalp, but, um, that only works for people with, you know, certain hair thickness and type. Um, so. That, um, you know, that definitely created a huge bias in just the people that we could collect data from.

Interviewee

Um, uh, even like, yeah, we like things like body temperature. could affect, like we would measure, um, you know, blood volume in people's finger, blood volume in people's capillary beds. And if it was cold and their, you know, blood vessels were constricted, then we wouldn't get a good signal. Um, Yeah. I don't know.

Interviewee

I feel like there's a very, kind of a wide Yeah. Variety of factors, especially in, you know, physiological experiments.

Interviewer 1

I see. Thank you. Um, did you ever use external data? So public dataset, third party, e p i Web script data? Yeah. Yes. Okay. And what, what were some of the quality issues?

Interviewee

Um, I think identity management is a big one.

Interviewee

Um, Uh, you know, being able to map, um, You know, someone's identity based on, you know, some com like fuzzy combination of their like name, address, you know, contact information. Like, uh, so, and, and oftentimes we would actually, we might ha work with a third party vendor to acquire a data set, but then we

would also work with another third party vendor whose job it was to do the matching between, um, between individuals.

Interviewee

So yeah, I think identity matching definitely a big one. Another one is just like, um, you know, making sure that, um, uh, like the columns, well, like, I guess it depends on, on like making, like if, if we are going to take a third party data set and we're also gonna take a. You know, a data set that we have and we're gonna, maybe we're like train taking a model that was trained on a third party data set, and then we're going to try to do like transfer learning on a model that's gonna on our own data.

Interviewee

Uh, then it can be definitely painful to make sure that the, you know, the exact, you know, the features match exactly between, um, uh, you know, their, their data and our data. Oh. Another example is we, we actually worked with, I worked on a project that where we had some, uh, apple Watch users and other Fitbit users.

Interviewee

Um, and you would think that, you know, step counts and calories and heart rates would, would be very similar, but we would, we actually found like population level differences. Um, we, we did some, some, you know, Small scale experiments where we had people wear like a Fitbit on one arm and a Apple watch on another arm, and we would get different step counts.

Interviewee

So it was, um, definitely, you know, those, those would be like, you know, we could, couldn't really solve for those because um, you know, we just had to. Um, usually we just had to accept that there was, you know, some, some level of uncertainty because we couldn't, we couldn't replicate results across the same, you know, across these two devices that are supposed to be measuring the same thing.

Interviewee

Um, there was one other thing about third party that I forgot. Um, well maybe it'll come back to me. Perfect. It's super

Interviewer 1

interesting. Thank you. , have you ever measured the quality of your data and or tried to improve it?

Interviewee

Uh, yeah. We actually used a well,

Interviewee

Yeah, I would say like a couple techniques we've tried, um, are just like setting, like in the case of like the physiologically impossible values, like we might, like set like rule-based constraints. Like, you know, someone can't be more than a hundred years old, someone can't be less than zero years old.

Interviewee

Like, and then we just implement those rule-based constraints and throw out either, either throw out or clip data, um, depending on, you know, how we wanted. How we wanted to account for it. Um, so that's one technique. Um, I mean, we've done, we've done like data quality monitoring where we set up like, um, you know, basically have dashboards that, you know, update daily, that we can look at, you know, um, the like central tendencies and variants of, of incoming data and see if there's.

Interviewee

Or end volume and see if there's any changes day over day. And if, if we do identify a change, then like reach out to the data engineering team and, and in charge of that ingestion process and try to, you know, track that down. So that's kind of like, not necessarily improvement, but just like monitoring and, and early kind of warning of, of data quality issues.

Interviewee

Um, what else? Um, oh, and then I, we, we have worked with, uh, Data robot vendor, they, they have like, this is more for like feature, uh, for predictive modeling. But like, um, if you train a model on a, on a training set of data, um, you would then, when you actually score that model on subsequent, you know, uh, feature sets, you would wanna make sure that those feature sets are like, reasonably similar to the, the data that you trained on.

Interviewee

Otherwise, um, the model could spit out, you know, different results because, you know, suddenly, Um, you know, there's, there's, there's been a drop in, like, the number of people reporting data or, you know, there's some shift in the central tendency. So, um, I forget exactly what measures they used, but it's like data drift essentially, which is just telling you when, um, when the features of your model don't resemble the training data anymore.

Interviewee

Yes.

Interviewer 1

And, uh, thank you. And what, what were the tools you used to, um, monitor your data?

Interviewee

Uh, yeah, I would say everything from like custom dashboards, like, um, well we, we were like working in Google Cloud, so we would use like data studio to create dashboards. Um, we've also just like created like Chrome jobs and Python, um, and, and generated visualizations that we just send over email.

Interviewee

So like purely custom processes there. Um, yeah, DataRobot would be another one. Um, And then I, uh, that, yeah, that's all I've viewed so far. I, I think we're, next year we'll start exploring, um, vertex ai. I think they have some built in tools for data, data quality monitoring.

Interviewer 1

Okay, great. I'll look into that. Um, is there any other data quality should we missed that you consider relevant?

Interviewee

Um, oh, actually, I just, I mean, I think like maybe extending on the thing I just said about data drift, like, you know, uh, we, I, I've been working in healthcare and the Covid 19 pandemic obviously had a huge impact on the way people, um, utilize healthcare. Um, and, you know, some of those changes were acute.

Interviewee

Like we saw like huge drop in the amount of people going to the doctor, like in the month of, you know, April of 2020. But, but then like persistent changes, like suddenly we see a, you know, sustained growth in people using telemedicine services because, or like, um, refilling their prescriptions by mail or something like that.

Interviewee

So I think like, there are these big, yeah, I mean, COVID is, you know, probably in a very extreme example of like a big world event that kind of fundamentally changes the landscape of, of the data that we use. Um, and. It makes it like, it makes all the sort of models that have been trained on pre covid data less reliable because the, the actual, like patterns changed.

Interviewee

Um, and then, and it's, we had this weird time where it's like, okay, we know the data's changed, but we don't have enough data to, to train new models because we, it hasn't been that long. It's, you know, um, so. We kind of just have to make do with what we have until, until enough time goes by that we can start retraining, you know, these, these, uh, models.

Interviewee

So anyway, I'm not just, just a very extreme example.

Interviewer 1

Yeah, that's, that's a good point. Thank you. Um, so I'll move on to model evaluation. So how do you evaluate the quality of your model? And as a reminder, um, quality is not only defined by the performance of the model mm-hmm. , but you also have other respect such as explainability, uh, fairness, robustness, scalability, uh, many more.

Interviewee

Yeah. Yep. Um, yeah, so I think. Yeah, I think actually sim so similarly, we, we did use like DataRobot to monitor model performance, but yeah, that, that was specifically looking at the model, um, performance in terms of like, you know, accuracy, precision recall, and um, you know, our, our sort of like model evaluation

metrics.

Interviewee

Um, in terms of fairness, we built, uh, an internal package that was based on, uh, I think it's ai, AI fairness 360, like has a set of like recommended metrics for evaluating, um, systemic bias for, uh, un like underprivileged populations. Um, so we, we implemented some of those, um, that was more, that was more for like checking models.

Interviewee

After they'd been trained, but not necessarily like, and, and then I think those checks get repeated on a yearly basis. So it's not like a continuous monitoring, it's more, you know. Yeah. Um, it was more of a, like one time when the model was changed and then again, sort of annually. Um, what else? Uh, yeah, I think we've done less around.

Interviewee

Runtime, um, monitoring. Um,

Interviewee

yeah, actually that's, that's interesting. Cause we, it's like we know when we, you know, train a model how long it's gonna take to score it. Um, but I actually, we haven't, I don't have any good examples of like, You know, monitoring how long the scoring process or how it like, you know, resource and sense of the scoring process is after the model's been trained.

Interviewee

Um,

Interviewee

and then what else?

Interviewee

Yeah. And then I think, uh, and similarly like, um, you know, I was mentioning Vertex AI for data quality, they also have some built in model performance monitoring. So we're gonna start leveraging those next year, uh, as well. Very great.

Interviewer 1

Um, yeah, and just, just I'm curious for fairness, uh, you use AI F 360, was it for c for your current company or your startup?

Interviewee

Uh, this was at Company X.

Interviewer 1

Okay. And what was the issue? Which group did you

Interviewee

consider? Um, we looked at, we, so we, we looked specifically for, um, uh, whether we were offering more opportunities to men. Um, Like, I think people between the ages of 20 and, and 50 versus younger or older age groups. Um, and then we looked, uh, yeah, we looked at to see if we were offering like more services to white, um, or people who, uh, responded as white in our ethnicity surveys versus people who responded as non-white.

Interviewee

Um, where those were the three buckets that we looked at. Okay. And what was the applic. Um, a variety of applications. So basically, um,

Interviewee

the like, uh, basically healthcare education programs, I would, I would say is probably the biggest one. So like, you know, letters, emails, and, you know, text messages that, um, would educate people about their healthcare options and, um, help them. Uh, kind of optimize their, their medical benefits, um, broadly speaking,

Interviewee

Okay.

Interviewer 1

Yeah, understand. Thank you. Um, have you ever accessed a quality of ML model prediction with the user of your system?

Interviewee

Hmm. Um, I'm trying to, maybe at the startup that I worked at, cuz we had a much smaller population. Um,

Interviewee

yeah, it, yeah, actually that, that is, so we had a much, it was, we only had a population of like less than a hundred people, um, in our, in our trial. Um, and uh, yeah, we. , we would give people like predictions of like, of certain cardiovascular metrics, um, and yeah, collect feedback from them if, if the, if the, um, if those, you know, predictions did not, you know, sound right.

Interviewee

Um, so yes, I haven't, we haven't done it at Company X, um, specifically, but this, but at the startup we. Okay. I see.

Interviewer 1

And what were some of the issue the user mentioned?

Interviewee

Oh, just mostly extreme values, I would say. Or either extreme values or values that never changed, um, over time. Okay. So

Interviewer 1

ml performance issues Model.

Interviewer 1

Performance

Interviewee

issue. Okay. Model performance issues.

Interviewer 1

Um, is there any other quality issue that we missed that you consider relevant about model evaluat?

Interviewee

Um, not off the top of my head, no. Perfect. Thank you.

Interviewer 1

Um, what are some of the challenge you encounter when you deploy model? Mm-hmm. . Um, actually machine learning software system. Yeah.

Interviewee

Yeah. Um, so I think. Currently, you know, at Company Y, like we, uh, we tend to like the, the data science team that I work on doesn't usually is in charge of like, training models and, you know, evaluating experiments, but not putting models into production.

Interviewee

So I think the biggest pain point is the, the actual handoff from the developers of the model to the engineering teams who are responsible for implementing the models. Um, and I think that there's a sort of a, um, Uh, I think that, you know, it's like there, there that knowledge transfer that happens where it's like the people who train the model have all this domain expertise, but no idea how the production, um, system works.

Interviewee

And the engineers know all about how to put things into production, but less about the actual domain, um, of, of what the model's trying to score. Um, so I think, yeah, I think that can. Somewhat of a painful process and

Interviewer 1

which, what, what, what are the consequences of, uh, this teko between, uh, software engineer and data?

Interviewer 1

Like one issue in particular?

Interviewee

Yeah. I think, um, when data quality issues arise, um, and there're being, well, I guess may, data engineering, data engineers who don't, who aren't familiar with the. Uh, like the, the sub, the domain of the experiment or the model, um, might not catch data quality issues because they're not, um, you know, they're, they're not thinking a about the specific, you know, use case maybe, um, and then the data science team who, who train the model doesn't have access to the same like monitoring, like software or tools or, or whatever's being used.

Interviewee

So they just like, so it is kind of, I, I think the biggest thing is like data or quality issues, um, with model performance go unnoticed because like the engineers don't necessarily know what they're looking for and the data scientists don't have access to that. Uh, those are those like production platforms.

Interviewee

Yeah,

Interviewer 1

Ok I see, thanks. Um, did you ever add a model that performed well locally but poorly once deployed?

Interviewee

Um,

Interviewee

I mean, I think I've had models that, um, degraded in performance over time. Um, once deploy. I can't think of any examples of things that, um, failed, you know, like that where, where the, the difference was like immediately apparent. Um, uh,

Interviewee

trying to think. I feel like that's probably happened. I mean, Right. And I think the biggest culprit there was, you know, be just because like, um, probably just like insufficient, uh, , um, Uh, use of like, you know, train or testing versus validation data, right? Like I think sometimes people don't actually create a separate test and validation set.

Interviewee

They just, you know, look at how their model performs of the validation set and, you know, constantly re randomize and, you know, so there is, they don't have like a true test set. Um, and then when the model does deploy, You know, the performance is worse, but I, and that's almost, yeah, that's expected just because, um, you know, it's, it's real world data, um, versus data that, yeah.

Interviewee

Yeah. So I, I think, oh, actually we did have a couple, uh, I can think of one particular instance where we had, um, we actually identified a data quality issue. Um, because we didn't, we didn't realize that there was like leakage in our training, uh, data leakage in our training process, um, until after we put a model into production.

Interviewee

So yes, I do have an an example of that. Okay. Thank

Interviewer 1

you. Um, issue with, with data during the maintenance of software system.

Interviewee

Um, Yes. Uh, yeah, we actually had, like, we've, and, and one of our applications, um, I think it was just, it was, yeah, actually it was related to a rollout of an app version, um, caused like a huge loss in certain types of data that were being, you know, used by our models. Um, and, uh, so it just, it just like the number of people who were being scored by those models just.

Interviewee

Dropped significantly, um, because there was a change in the data ingestion process. Um, okay. I see.

Interviewer 1

And now, do you have any mechanism to prevent this kind of issue in the future? I think change as, yeah.

Interviewee

Go for it. Yeah, that, that's when we started, um, implementing like more, uh, of these like monitoring dashboards.

Interviewee

So at least we could, we could iden, like identify on a day-to-day basis when there were issues with, with data ingestion, and then how those issues, uh, affected like the downstream, um, model scoring.

Interviewer 1

Okay. Perfect. Thank you. Uh,

Interviewer 1

Is there any other issue regarding the maintenance or the deployment of machine learning software system you'd like to discuss?

Interviewee

Hmm. Yeah. Well, I mean, I think one, one was just like , um, I don't know. We, I work in a very large company and there are a lot of different teams developing models and oftentimes we have no, I like, we just have, first of all, we don't have a, a robust inventory of all the models that are in production.

Interviewee

Um, so it's almost, yeah, it's like people end up training brand new models. You. because they didn't even know that the models that, you know, that there are models that already exist in production that are doing the thing that they wanna do, but because there's just like a gap in, there's not like, there wasn't like a good inventory system.

Interviewee

Um, That was another reason to that motivated us to, to start, you know, to start exploring Vertex ai cuz they have like a model registry. So yeah, not having a model registry leads to a lot of duplication of, of models that doesn't need to exist. Okay. Interesting. Thanks.

Interviewer 1

Um, . Now I'll mention some quality aspect, and you, you, if you ever have, if you ever add an experience, if you ever add an experience in which you had an issue with the quality aspect, uh, feel, feel free to mention it.

Interviewer 1

Yeah. Okay. Uh, so fairness, you mentioned it already. Uh, robustness, explainability, scalability, privacy, and

Interviewee

security. Um, definitely explainability was one. Um, and, uh, so this, this, we had wanted to use, uh, I forget what like, um, uh, ba basically we had a model that was affecting a eligibility for a healthcare offering.

Interviewee

Um, and by regulation like that healthcare offering had to be, um, Had to be well documented in like the, the, like, you know, the handbook of, for our, for our members. Um, and so we ran into an issue where, you know, the model that we were using, uh, wasn't like sufficiently explainable to users. So we had to sort of downgrade to a more explainable model for compliance reasons.

Interviewee

Um, be because that model was used to offer. Um, you know, to offer a service to people, uh, who we predicted needed the service. Um, and then I think there was another thing where it's like we had to, then we had to then make that service available by request, regardless of what the model said. Um, so it was definitely, you know, it was an interesting like, um, issue to come across because, you know, we were, we were only thinking about, you know, how well our model performed.

Interviewee

You know, we, we were encountering a regulation that made, made us prioritize the explainability of the model and then also made us, you know, offer this program to people regardless of what the model said. . So,

Interviewer 1

yeah. Yeah. Good. Interesting. Sorry. Thanks. Uh, so yeah, we have two more questions and we're finished.

Interviewer 1

Okay. Uh, so in your, in your opinion, what is the most pressing quality issue? Researchers should try to,

Interviewee

I think, uh, yeah, I'll just go with the first thing that popped into my head was, is just, um, reproducibility. Um, cuz I think, uh, you know, oftentimes, we'll we leverage, you know, I, I guess it, it kind of feels like there's not, um, There's not a reason why we can't reproduce things. Um, particularly things that are published in academic journals.

Interviewee

Like they're, like, you know, publications should make the data available and the code that was used to analyze the data, um, uh, so that people can reproduce those results. And, um, yeah, I, I, I think that.

Interviewee

Yeah, I don't know that, that, that feels the most relevant to me. Um, okay. Granted, I dunno.

Interviewer 1

Yeah. Reproducibility only of paper, but, but not of experiment you did, um, as a professional, not only for Right. That's a Okay.

Interviewee

Uh, I think, I think it is that also there's also. Implications in industry, although, and, and I think maybe there the reproducibility needs to occur, like within the industry, maybe.

Interviewee

Like I, I understand people will always want to protect their trade secrets, but like, yeah, I would say even, you know, even within our company, we've, we've struggled to reproduce results like, you know, especially. You know, someone builds a model and then leaves the company and, you know, the model's still running in

production, but no one knows how it was, no one was, no one who was ar was around when the model was trained.

Interviewee

Um, so yeah, I, I think re I think that, that, that applies both to like academia, but also to like to industry as well. Thank you. And I have a follow up question. Yeah. That with. Make any problem in your real values, in your procedures, in your, in the, the way that you try to deploy the train model. Um, yeah. So two things, uh, kind of, so one is, so in health insurance, um, Often the value of a model is, is perspective, right?

Interviewee

Like the model of an in an intervention is perspective. Um, so, you know, let's say I have a model that can, that can predict, you know, who's gonna have a heart attack in the next, you know, 12 months. And, um, we're gonna target those people with, you know, education about pharmacological agents that could help prevent those heart attacks.

Interviewee

And we're gonna say like, every heart attack we prevent is gonna be worth, you know, $10,000 or a hundred thousand dollars or whatever, like, um, we, we won't be able to measure that until, I don't know, 12, at least 12 months later, probably longer. Um, so when we, when we launched this, you know, project, we have this like prospective value and then.

Interviewee

We have to measure it later. Um, and, uh, when, when that, when, like, those programs don't like meet their, the expectations that were set when they were first launched, like oftentimes, like, you know, people wanna know why. Um, and that can be really difficult when the people that, you know, initially created those models are no longer there

Interviewee

So that's happened more than once. Um, uh, so that, that's one piece of it. I think the other piece of it is, Yeah, like if there are changes in the data because of Covid or because of some, you know, data or become app update or, or whatever, you know, changed to your ingestion process, like, that means you probably have to

retrain the model.

Interviewee

And if the training process wasn't well documented, then um, you, you have to start from scratch.

Interviewee

Yep. I think.

Interviewer 1

Yeah. Thank you. All right. Yeah. Um, so do you have any other comment about the quality of machine

learning

Interviewee

software system, um,

Interviewee

thinking? I don't think so.

Interviewee

Yeah, I think, um, Yeah. One other, one other thing that I feel like everyone has talked about in our company,

but I haven't seen implemented, like in a robust way is like, um, like indi, like, you know how like, uh, shap

values give, like individualized, you know, uh, sort of explanations for why, you know, specific people were

targeted by a model.

Interviewee

Like, um, I feel like everyone talks about how that would be like a handy feature to have like available by

default, you know, especially. If we're offer, if we're using it to offer people, you know, products or services.

Um, and, uh, yeah, I don't know. I, I guess that's something that I've heard talked about a lot, but I've never

actually seen implemented at scale.

Interviewer 1

I see. All right. So, uh, it was pretty interesting to have you today, Interviewee, everyday pleasure. And I think what you, you share with us today will be useful for a study. And so

Interviewee

thank you for being. Glad to hear it. Yeah. Um, yeah, great talking with you as well.

Interviewer 1

Yeah. Thank you. And am, am I mistaken or is the master sword behind you?

Interviewee

Uh, from link, that is the master sword. Yeah. . It's foam though, not metal, but yeah. Good eye. Okay,

Interviewer 1

great. All right. Thank

Interviewee

you. Yeah. Pay attention to everything. Take care. Happy holidays.

Interviewer 1

Thanks. Same, same to you. Have a

Interviewee

good Bye. Bye.

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