

Interview 33 - PO

Interviewer 1

All right, so can you give us a little bit of information about yourself and how many years of experience you have in advance learning or in general?

Interviewee

Okay, so, uh, I stay right now at City X ne uh, greater City X area, um, uh, to be precise. Um, uh, I also work for this company called Solidaridad Network. Uh, it's based in the Netherlands and we are primarily focused on ag agritech and, uh, with machine learning, I've, I've been with the company.

Interviewee

after college, say it's around two years and seven months now. And, uh, I started, uh, with an ML role, say I think, uh, one and a half years ago. So around two years of experience in ml. So, uh, before, prior to this company, I has to work with this, uh, education startup where we build courses for ML and data science foundations, you know, and, um, in this company, I mean, I think my first project was on junior role was to build a chatbot, uh, which engaged with farmers, uh, uh, that I think that was the first role.

Interviewee

And, and the best, uh, project that I've had with this company is, uh, we work with a lot of. Tea manufacturing stakeholders, you know, tea factories, uh, tea processing units, tea collection units, farms and all that. And there was a huge, uh, problem, uh, impending Was that, uh, the quality of tea leaves, not the tea powder.

Interviewee

The tea powder was very, uh, easy to sort cuz it was in a close loop in a. But the tea leaves just came from all over the world, and they were just agents bringing them from, uh, from small farmers to big farmers. And, and there was no particular way to, uh, you know, kind of assess tea quality, uh, the leaf, uh, quality.

Interviewee

And we built a system called Bellas, which we have a patent. And, uh, patent just got published I think a few months ago. Uh, so we used a. I mean, uh, uh, uh, a deep learning model for it. And, uh, we came up with, uh, around 89%. That was our last validated, uh, accuracy of the model. Uh, this was, uh, uh, validated by the institute, the Research Institute, research Institute of Country X.

Interviewee

Uh, so they took our machine and the model, and they tested it. and they've come up with this, and it's approved by, uh, the research foundation, which is run by the government to use NT factories. And that is my ex, my mean, my best experience in ml. So yeah, that's, and I, I shifted from Country X to Country Y a month ago, and I've been working remotely from here.

Interviewee

And why I shifted here is, uh, I've been, I'm trying to work and pursue a postgraduate certificate in project management. Uh, so yeah, that's, that's about me. Okay.

Interviewer 1

Uh, so to start off, can you, uh, well what are the main quality issues you have encountered with data model or

Interviewee

systems so far? Okay, so, um, Let me, uh, use a st, uh, project that I just spoke about as a best example.

Interviewee

So we started off with, um, just using a, um, say a statistical model. Uh, we, um, we worked with, we thought, we thought we could build it initially using just, uh, uh, the pixel properties, the size of the leaf. And those kind of things. And, uh, after our final test, it, it, it didn't hit anyway, you know, the, the, the quality was too low and the accuracy was too low with, uh, with using it.

Interviewee

So, I mean, that was the first step. Then we had to eliminate that, and then we moved to a deep learning model and deep planning model. We started with fa fast, c n n, and. Again, the problem was, uh, the type of, uh, input data or, or the way we did it was, was wrong. We, we, we took it whole branch of, of, of the Tealeaf, and, and we tried, uh, you know, comparing that with other, uh, categories of tea, but that didn't work out as well.

Interviewee

I mean, it's, it's not that it didn't work out, it, the accuracy was way below what is acceptable. And then, uh, I think we finally switched to a mosque, R C N N. Um, So that really helped. But in that as well, we had quality issues. I mean, the first data collection went on with, uh, our field staff just going and clicking pictures, uh, o of these leaves on the field.

Interviewee

And, uh, when we put this to test and when we put this to validation in the factories, uh, the accuracy was low. But, uh, I. We, we've done our best coming from, uh, a statistical model all the way to Mascar, cnn, but, uh, we just couldn't figure out what, what it was. And then, then we figured out that, uh, if we could take, uh, samples from the machine itself rather than the field, uh, we got at least, uh, 12 to 13% improvement on our accuracy.

Interviewee

That is, I mean, I think that is my experience on, uh, dataset collection and, you know, quality management and I'm talking of. Say 20,000, 25,000 leaves. So it's a huge process and there's like 10, 12 people working just on data collection and uh, uh, you know, those kind of things. So, so, so that's, that's, that's my, uh, experience on it.

Interviewee

Okay. No. And there was no at prior dataset at all, anything at all? In, in the, on open source in the ma in, in the internet. Nothing. Nothing that has a Tealeaf dataset, so I think it was our first dataset, so, I mean, we had a village ground up, so it was, it was amazing to learn as well.

Interviewer 1

Nice. Uh, what did you meant by, um, uh, sorry.

Interviewer 1

Uh, I think you mentioned that you took sample from the machine instead of the fields. What did you mean?

Interviewee

Oh, yeah, let me correct that. Let lemme correct you on that. So, uh, so as I said, uh, the final implementation of this project was, uh, a small scanner like machine, uh, with, uh, a small MicroCon. We, we use our aspray pie inside.

Interviewee

So ba basically the PIs, uh, took these images from a wide, uh, wide angle, a 10 80 p camera. and send it to our servers. We had a flask server running, which, which could take these, uh, uh, images and then, uh, uh, you know, this, this mask model was running, and then it, it'll give back the results, the, uh, adjacent response of all the results.

Interviewee

So, um, I'm in, in, in that whole thing. So that is a machine that I'm talking about. So first we had just cameras say like, do you sell our cameras? We went around the field and took focused images of the leaves. Uh, later what we did is we, we took images, uh, directly from this camera of the, of the machine. I mean, okay.

Interviewee

All the input image was just like that. And then what we did is from that and from that huge, uh, Or you say that huge frame. Then we picked out the images from that. We used Label Me, I think, I think we used Label Me and then we, uh, do the, uh, labels and that's how we took the, uh, that's how we made the second version of the data set compared to the first one.

Interviewee

So the first one was more focused on a faster c n n model, so we didn't have to necessarily draw the boundaries, it was just bounding boxes. So that's why we went with just taking. Photographs of individual leaves, whereas in the second one, we went to the ma uh, Moscow, cnn. So we had to draw the boundaries.

Interviewee

So we just got huge, uh, the whole f uh, you know, the whole set in one frame and then we drew the individual leaves from it. So that considerably improved, uh, the model performance. So that's what I meant. Okay. I understand. Thank

Interviewer 1

you. Um, did you ever use, uh, Sorry, sorry. Uh, did you ever use external data? So public data sets, uh, third party API and web script.

Interviewee

Uh, so as I said in, in this, uh, I have done in, in other projects, uh, you know, site projects that we worked on. But in this project, particularly as I said there, there was no data outside. Uh, we, we had, uh, after we've implemented the first phase with this project, in the second phase, we had this idea of implementing, uh, a disease recognition model, uh, where we un we, uh, recognize patches of disease on.

Interviewee

And estimate how much, how disease the leaf is. So in that case, to just start off with, to do, do a proof of concept, we found a small data set of leafs, uh, uh, of t leaves. And yes, we did use that and we got some results, but we didn't push it forward because there was, uh, at, at that moment there was no market requirement for that particular model and that particular feature.

Interviewee

So we haven't implemented it yet. But yes, I did use that for that. Okay.

Interviewer 1

And was there, what were the main, sorry, what were the main issues with, uh, this

Interviewee

type of data? Um, this, uh, I can I trade it back to what I just said earlier? So these images were again, uh, across of, uh, all the da, all the, uh, type of captures.

Interviewee

I mean, some of it just, uh, photographs, uh, somewhere, you know, a whole. The whole bunch, a whole branch or, or sometimes it'll just over, uh, you know, the white balance was not proper. Uh, or it was some, sometimes it's too dark. Uh, it was not, uh, to be, uh, uh, you know, to give a whole picture. It, it was not. Uh, it was not regular.

Interviewee

The data was not regular. It was very off. And, uh, I don't think we could have made a enterprise model or we could make something concrete with that. Uh, but probably something like a P O C just to prove that, okay, we can do this as well. Uh, something to show our management that is all the, the dataset could do.

Interviewee

Uh, what I believe is, uh, if, if we are looking at making a concrete enterprise model, we, we have to build that data set ourselves. And keep it as, uh, as similar as possible. Yeah. Thank you. Yeah. Um,

Interviewer 1

have you ever measured the quality of your data and or tried to improve

Interviewee

it? Uh, yes. So, uh, how, how we started off our data collection was that, uh, so, uh, I, we, I don't know if you know this, but then in Country X, uh, different parts of Country X, even though it's just one country, different parts of the country has say five or six varieties of tea.

Interviewee

And, uh, when the model performed well in one variety of tea and when it was taken to a northern part of Country X, The leaves were just, uh, the same category of leaf looked like another category, you know, things like that. So it, the accuracy was very low. So we had a lot of false positives then. So, um, our data collection

majorly was done in factories.

Interviewee

So when these, uh, leaves come in, uh, when they come in to dump their leaves, that's when we take a part of those leaves and then we do the data collection. So when we saw that, uh, this was happening, that, you know, different varieties of tea, The, the accuracy was becoming low. We had more false positives, so we started implementing this data collection.

Interviewee

We started reducing the data collection, uh, say from a single factory, and then we started to distribute it, you know, take it from different places, even though it's less a number from different places. And we did see, uh, a small increase in performance while doing that. Okay, I

Interviewer 1

understand. Thank you. Yeah.

Interviewer 1

Um, is there any other data quality should we miss that you consider?

Interviewee

So, um, what I would say is, uh, data augmentation is a huge thing. If, if, if you want to like build, um, say strong models, uh, It, it might look irrelevant when, uh, say in, in our case, uh, the machine, uh, everything is fixed, right? The lighting is fixed, the size of the image is fixed.

Interviewee

Uh, and the type of leaf, it's a T leaf. We always know it's a T leaf and it's not gonna be anything else. But even in that case, I could see data just augmenting, say the position and wide. Significantly, again, improved our da uh, our, our, our model accuracy. So there, there are times when these, uh, uh, the ma in this machine, the t leaves are backlighted, meaning the light.

Interviewee

There's a light coming from the back, and that's when it project projects, uh, the properties of the, of the leaf, right? So this can be different or the, you know, based on the exposure time of the. Uh, the white balance might change here and there, but then, uh, we didn't take this into consideration earlier in the production, but later in production, we thought, why, why not try that?

Interviewee

You know, why not try some data augmentation with white balance? And we were surprised to see that it, it helped us so much. It, uh, a lot of false positives, for example. The outliers, the layers, the leaves that were almost in the outer section of the frame were just too exposed. Uh, and we couldn't kind of, you know, most of the time it was not even detected on the model.

Interviewee

But then when we tried doing this, the detections increase, you know, even though outliers started to get detected. So, yeah, that, that is one, one thing I could add. Mm-hmm. . So data

Interviewer 1

augmentation plays a huge part in.

Interviewee

Yeah. Yep. Perfect.

Interviewer 1

Um, how do you evaluate the quality of your model? And as a reminder, quality is not only defined by the ML performance, so accuracy front score, uh, but there's also other aspects such as explainability, scalability, efficiency.

Interviewee

Uh, at least in, in my case, uh, there were two, uh, overriding factors. Uh, one, uh, we are talking about replacing a particular, uh, system that has been in place for almost 300 years. Uh, three centuries. This, this way oft, uh, quality assessment was being done and, and, uh, a company like us just comes in and, and we we're trying to change the whole thing.

Interviewee

So, uh, what we had to do in that case is that we had to. You know, compete with the manual process. So what we did is, in our cases, uh, we had the manual person, there was a person who was doing this earlier. It was just that the machine is doing it now. So we made a monthly, uh, uh, quality assessment. Uh, Schedule.

Interviewee

You know, I, I, I have those sheets. I can share those data with you. So we compared what, uh, when the batch of leaves came in and these were separate, you know, they, they were not in the same room. They're not biased at all. Uh, this guy doesn't know the, what, uh, the guy who's handling the machine doesn't know, uh, the, uh, about the leaves.

Interviewee

The manual guy is doing, but it's the same leaf. You know, the leaves are put in a basket. He does it first, and then it goes to the machine, and then the machine does the evaluation as well. And we compare these results for three or four months. And then, uh, we came at, uh, Six to 7%. We came that close when we came that close.

Interviewee

Then we, uh, I mean, as you, as you rightly said, ML performance is a big thing. That's, but, uh, when, when, when you implement and what is the end goal of an ML project to be implemented in a real, real world situation right there, it. Uh, the only factor that overrides everything is how does it compare with something that was manually done already, how well does it already do?

Interviewee

So in that case, when it came up to 7%, that's when we commercially started, uh, selling them. That is number one. And number two was that, uh, we are talking of, uh, 40, 50 farmers standing in a queue waiting to get their leaves assessed. So the main prob the main, uh, A factor there is that the speed of the a p l response, you know, that was again, uh, a huge thing.

Interviewee

So our initial, uh, with all the complications and all the things that we did, all the free processing steps that we did, uh, before we, uh, turned it into the model, it took, uh, I think around 50 seconds, 50, uh, 45, 50 seconds for the whole response to come back. And then we reduced that to 25 seconds. We removed a lot of steps that we fed, irrelevant to the whole process.

Interviewee

and, uh, in our case, those are the two stakeholders. You know, how does it compare with the manual process and how fast can we give this response so that this becomes a profit system that we can implement in a factory?

Interviewer 1

Why were there, why were the step irrelevant for, uh, your products? Sorry? Why were the, the, sorry.

Interviewer 1

You mentioned in your pipeline you removed some. Yeah. Yeah. What, why?

Interviewee

Yeah. Uh, for example, um, there was one there, there was one particular step where.

Interviewee

For example, uh, in the pre-processing steps, uh, for the T leaves, there was, uh, there was a step where we, uh, cropped the image and, uh, gave a fixed image. I mean, um, That that step really did not when, when, when we, uh, I mean these things were done when we were really out of time and we just did it cuz we had to do it.

Interviewee

And then we had, when we had time to like, you know, kind of review back how the production is. And uh, that time when we saw that cropping really did not, uh, make a huge impact on the performance with even the model without cropping, it didn't, uh, was fine. I mean, it worked fine, but then that took at least.

Interviewee

Four seconds or five seconds of our processing times. That is one. And, uh, oh, yeah. Let me come up with the most important one that we, that we solve is that, uh, I just remember this now. So earlier what we did is that, uh, the scanner, the camera that we used was a normal camera. And, uh, we, uh, we kind of moved it on a, on a, on a track and it took like four or five images and then we did a pan shot.

Interviewee

We stitched these images together and that's when we, then, that's then when it, when it went to the model. So this whole thing took a lot of. Uh, and why, uh, would we just use a wide angle camera? We, we thought there would be bending on the sides and, uh, you know, the bending would again, uh, you, you see that wide angles can bend on the sides, right?

Interviewee

And the, and the bending would affect the moral performance. But then, uh, this started to counteract, you know, , it was just taking so much time. The whole thing of going and coming, it takes around 20 seconds and then the stitching takes another five seconds and then sending it the model to work. It was taking a lot of time.

Interviewee

Then we just switched to a wide angle model and then, uh, we fed in more, uh, uh, uh, on the input images we fed in more of these WideAngle cameras and, uh, annotate, we annotated more of these wide angle images. It worked at the end and, and we removed that whole track system and, you know, we, we provided a second version of the machine where the track was not there.

Interviewee

Uh, there was, there was nothing. It was just a wide angle camera, no noise, no nothing, no moving parts, just one shot. And I think we saved, that's when we came from that 40, 50 seconds to like 20 seconds, at least 20, 25 seconds. And then, as I said, the cropping and, uh, I don't know, a few other small image processing steps when we could remove that, we could shave off like two seconds, three seconds here and there.

Interviewee

And it didn't considerably, uh, affect our ML performance, our model performance. I see, I see. Thank

Interviewer 1

you. Um, what are the challenges? Challenges we have encountered during the deployment and the maintenance of a machine learning

Interviewee

software? . Uh, so I mean, uh, I thi this is, uh, technically there were a lot of, uh, um, uh, uh, hurdles we had.

Interviewee

Number one hurdle is that, uh, we had to sell this and we had to scale this, uh, pro project, right? But, uh, We couldn't get the hardware required to run this same mask model into, into that machine, so it'll just blow up our costs. You know, we, we had to send this to our server and get it done there. But the problem is that, uh, these t factories are just located in the most remote areas and.

Interviewee

I'm not even talking a 4g. Getting two G is, is, is, is a task, you know, so, uh, packaging this whole image and sending it and receiving the response and sending it back, it, it was a huge thing. So, uh, we had to cut down our, uh, Resolution we had to compress the image. Once we compress the images, uh, you know, we had to kind of find that balance between, uh, these, these, this much data, these many pixels you put in, and this much information, uh, versus the time for the whole thing, for the whole response.

Interviewee

So we had to balance that. And then we, I think we came up with say, 30% compression on our input data. Uh, that is number one. And number two is, , even though we make, uh, other software projects, uh, say, uh, say a data entry system, you know, some, just a normal front and backend data entry system is making someone's life easier.

Interviewee

You know, so everyone's, uh, looking goodo look, looks at it with, with a good face. But the problem with ML systems is that ML system replaces a human. You know, that is a huge thing. I mean, when we implemented this, The tea guy who, who was, you know, the main boss. He was the main dude there. You know, he decides the quality of the tea.

Interviewee

He's a main biased guy, you know, he, you, you, you piss him off. You, you lose your quality. You, you lose your price. And now there's a machine that comes in, an ML model that's come in and it's completely unbiased. Uh, and that does the, and that runs the whole show. You know, that is the number one hurdle. ML systems replaces the roles and to convince a management or to convince, uh, an a community about how this might be better is a huge task.

Interviewee

And that's exactly why we had to run this manual versus machine mo, uh, test for. Three or four months to prove to the higher management saying that how unbiased and how better we can, because this guy took around 15 minutes to do, uh, uh, a particular set, and we took 20, 25 seconds to just assess a, a set of leaves.

Interviewee

So yeah, that, that is, uh, two biggest hurdles that we had. I see. I

Interviewer 1

see. Thank you. Um, so to reduce latency in your call, you compress the. Yeah. Before sending it to your server. Yeah. And um, so, and so the model was predicting, uh, was on a server not, not on a respiratory tie, no. In the fields or, okay. You simply

Interviewee

couldn't run there.

Interviewee

Yeah. Okay, perfect. Thank you. We even tried, uh, experimenting with, I don't know if you know about this, uh, uh, jet from N Video. Uh, we tried the two GB A G B models, but none of them just ran well. I just took so much time and it just didn't make sense to run it there.

Interviewer 1

I see. Perfect. Okay. Um, I'm gonna name a few quality aspects and you tell me if you ever, uh, experience have any issue with one of these.

Interviewer 1

Okay. Uh, and if yes, well share with me please. Your, your, sorry. Uh, so fairness, robustness, explainability, scalability, data privacy, and, um, model security.

Interviewee

Um, model security is, uh, I think the vendor's problem. My problem, uh, robustness and, uh, data security. That is it. That is number one. You know, and we are talking of, uh, We are talking of implementing ML models for people, right?

Interviewee

Uh, for, for companies. And you just don't have one vendor. You have, say, 10 vendors and one vendor doesn't want, uh, his data, uh, or his, uh, way of qu uh, you know, measuring quality, uh, uh, to be seen by another vendor. So that's a huge thing that, that, that is number one, I would say. And second was robustness.

Interviewee

And third scalability. You can make a very fancy model that, you know, detects everything in this whole world. But, uh, if it doesn't scale, if it doesn't do things on time, uh, and if, uh, if it, uh, keeps failing, say after four or five, uh, a p responses. Doesn't make sense. Right. So I think, uh, sta number one would be data privacy.

Interviewee

Second would be robustness. And third with scalability. Mm-hmm. .

Interviewer 1

And how do you ensure data privacy in your case?

Interviewee

So in our case, we had, uh, end-to-end encryption. And the models and the data trained were, uh, discreet for each vendor. Uh, even though that, uh, reduced, we had an initial model, uh, that that could do, uh, the basic, uh, categorization of quality.

Interviewee

And after that to improve the model, we further took data from only that. and we made sure to inform that person that this data and the model accuracy will only depend on the data that you give me. It, it might be something different in a, on, on a different vendor. It may, might be different with another person, but it'll be only with you.

Interviewee

So, uh, we kind of built that trust with that person saying that, uh, you know, you have your access to this model, you have access to this data, and not even us, uh, have access to this beyond the f the basic model or the basic version of that. Okay. I see.

Interviewer 1

Perfect. Um, and I have two last questions for you. Uh, in your opinion, what is the most pressing quality issue researchers should try to

Interviewee

solve?

Interviewee

Um,

Interviewee

I would say, uh, I would say we should work on, uh, On models that, uh, I, I, I think we, we, uh, I don't know how much it has to do with quality, uh, but I, I think we should make more, uh, uh, what do you say openly, uh, adaptable models that is, uh, say mask, uh, has. Very small, uh, uh, category where it can be applied or say fast.

last-question-answers

Interviewee

RC n n has a very small category where it can be applied and we'll, most of the times we just bend our problem around it to, to, you know, kind of fit it. So researchers should, should, uh, uh, come up with more reliant, self-reliant models where, uh, the data necessarily don't ha doesn't have to be, uh, tailor.

Interviewee

For that particular model where it can, you know, models that can take up anything and churn out, uh, better results. I would say researchers should probably work on that. Uh, you know, it should be able to take out any, uh, so I know that we have models that can take low quality images or, you know, very old images and make them look like new or, you know, all those kinds of things.

Interviewee

Yes. But then I'm saying we should work on enterprise models that can take, that can, you know, work on a large, uh, scale. Say worse data to good data, and then yet come out with, uh, results, uh, like the present models.

Interviewer 1

Yes. Sorry. Each time I have to click and it takes a bit of time, so No problem. Yeah. Sorry for my late response.

Interviewee

Just to add, uh, yeah, I, I, I'd like to add this another, uh, it just came on my mind and I thought I'll just put it up press, so, uh, you, you, you had put a point about, uh, increasing quality, uh, so I mentioned that, uh, response.

Interviewee

1, 1, 1 of our major parameter. So another thing that we did to increase our response time was that, uh, we moved our code from synchronous to asynchronous. Cause we found that a lot of time was just getting wasted in IO operations where we were waiting for the image to come in and then we wait for the image to go out and receive and all that.

Interviewee

So when we moved our, so in our case, we started off using flask, our API framework, and then we moved to Fast api and then we moved to an asynchronous code. And again, improved a lot of things. So that is another thing that we did to improve quality. Yeah.

Interviewer 1

Mm-hmm. . Yeah. And do you have any other comment about the quality of machine learning software system?

Interviewee

Uh, none. I think, uh, that's all I would've said. Uh, nothing much to add. Yeah.

Interviewer 1

All right. Well thanks a lot, uh, you for being there. We really appreciate and I think you mentioned a few good points that will be, uh, useful for our study. Oh,

Interviewee

okay. Yeah. Thank you so much. So.

Interviewer 1

Have a good weekend. You, you

Interviewee

too. Thank you, Gina.

Interviewee

Thank you

Interviewer 1

for your time. Thank you. Have a good Bye. Bye-Bye bye.

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