Transcript

00:00:00 Interviewer

So, I'm recording your session of your after your permission.

00:00:08 Interviewer

Yeah, I've been.

00:00:12 Interviewee

Free since last year.

00:00:18 Interviewee

I am based in Pakistan by the way. So, from fast and then I joined in the software industry. So I'm mostly working on full stack as a professionally. I've been working on full stack development.

00:00:32 Interviewee

AI has been something that I have been exploring academically. I would say. I mean in like.

00:00:39 Interviewee

I had couple of courses from Audacity and things like that, but I haven't had done any professional work. As I pointed out initially that most of my experience would be academic with respect to machine learning, but of course I mean I will try my best to help you where you like. Where can I be useful?

00:01:00 Interviewee

Uh, one question are are you from like Pakistan?

00:01:04 Interviewer

Yes, I am from Pakistan.

00:01:06 Interviewee

OK. So you are like the incident for masters? Or have you shifted to Sweden for a while now?

00:01:15 Interviewer

I I've been here for more than nearly three years. I'm working here.

00:01:19 Interviewee

OK.

00:01:21 Interviewer

But now I doing my studies part time.

00:01:25 Interviewee

Got it, got it.

00:01:27 Interviewee

That's great. Yeah, that's great.

00:01:30 Interviewee

So yeah. Well, we can move to the next step, I mean.

00:01:34 Interviewer

Yeah, we have. Do you have questionnaire often in front of you?

00:01:35 Interviewee

How do you want to go? Go ahead with that.

00:01:40 Interviewee

Not yet. I can open that.

00:03:19 Interviewee

The question I did you had sent me on mail, right?

00:03:26 Interviewee

I can't find it. Can you send it again, I mean?

00:03:40 Interviewer

Can I send you through LinkedIn or?

00:03:44 Interviewee

Yeah, you can send me on the mail. I'll.

00:04:03 Interviewer

If you follow the same thread.

00:04:04 Interviewer

Which I have submitted the zoom book.

00:04:07 Interviewee

OK, I'll.

00:04:33 Interviewee

So right now, are you working on ML or like your job is related to ML or something else?

00:04:43 Interviewer

No, my job is as a full stack developer, so mostly I went in the.

00:04:48

OK.

00:04:49 Interviewer

And then.

00:04:52 Interviewer

C#.

00:04:57 Interviewee

Jordan. OK.

00:05:00 Interviewer

Yeah, I send you the e-mail.

00:05:04 Interviewee

Just checking, sorry.

00:05:17 Interviewer

Yeah. So we will go to the questionnaire from top to bottom.

00:05:24 Interviewer

So if you have any confusion, you can ask me again, I will open up right now.

00:05:30 Interviewer

Just where is it here?

00:05:33 Interviewee

So yeah, I'll just answer this question and that's all right. I mean.

00:05:39 Interviewee

No, there's no need to fill anything, right?

00:05:42 Interviewer

No, I'm recording so I I can do analysis later on, yeah.

00:05:44 Interviewee

OK.

00:05:48 Interviewee

OH god got it. Got it. So sure.

00:05:53 Interviewer

First, the first question is can you please introduce yourself and describe your job role in this company?

00:06:01 Interviewee

So yeah, I've I'm full stack engineer, mostly working on front end and front end and back-end technologies in my professional career I have been mostly I have actually explored. I have been like for three years development developingthroughwith.net.

00:06:21 Interviewee

Technologies. And then I shifted to open source JavaScript react technologies. So right now in my company I am I'm the team lead with the six member team.

00:06:34 Interviewee

And I managed those like the team day today work and stuff with product man discuss different product related decision with product managers and stuff. So yeah that's that's the current role that I have in my company.

00:06:52 Interviewer

Since how many years you are working in this company?

00:06:56 Interviewee

Yeah. So from the in the latest job has been for I've been in this job for three years and before that I was another company QMD that's that was also three. Yes, so 3.5 and yeah, same is in here 3.5. So yeah.

00:07:13 Interviewer

Yeah. So you have how many total like 7?

00:07:17 Interviewee

Seven. Yeah, approximately 7 days.

00:07:18 Interviewer

OK, YOUR engines 3.5. And what about the current company?

00:07:22 Interviewee

3.7 I think, yeah.

00:07:27 Interviewer

So is it fine if you share your current company name?

00:07:32 Interviewee

Yeah, current company is “” based in Lahore. Yep.

00:07:34 Interviewer

I I.

00:07:40

OK.

00:07:41 Interviewer

So, have you published any thesis in the machine learning systems?

00:07:45 Interviewee

No. Yeah, as I pointed out, I've I've mostly just started exploring it recently since last year. So I'm just, I'm still on the in the phase of learning the technology.

00:07:59 Interviewee

So I have done a couple of nano degrees programmes but haven't published anything related to, you know, machine learning thesis stuff so.

00:08:10 Interviewer

OK, so have you worked any machine learning projects?

00:08:17 Interviewee

In the Machine learning project that I did were mostly related to that nanodegree programmes so.

00:08:25 Interviewee

I did a few. There was one sentiment analysis once that that would.

00:08:31 Interviewee

Like find the sentiment of in IMDb review that I mean any movie review on IMDb. And then there was this another project called there it was classification of flowers based on their like images.

00:08:48 Interviewee

So we, like, did an animal train a model to predict the name of?

00:08:54 Interviewee

A specific species of a flower based on our tester, our training data. So that was one. And then the final project that was mostly like the main caption.

00:09:09 Interviewee

Project that they offered was about Starbucks, where I where we did some data analysis to come up with a solution that could be used to.

00:09:19 Interviewee

Like get the best offers for customers for Starbucks customers. So yeah, there was an open data set, so we did that analysis and made a model, yeah.

00:09:35 Interviewer

Fine. So I think so question #4, can you please your experience the current your job current position?

00:09:46 Interviewer

Do you have lunch ready?

00:09:46 Interviewee

Yeah. So.

00:09:49 Interviewee

Yeah, so currently I'm mostly working on front end technology and there is some exposure of air stuff, but mostly on the discussion level, I do not.

00:10:00 Interviewee

Literally like look into the most AI stuff like model training, so professionally it's more like decision making stuff that we decide that what should be done and what should so.

00:10:14 Interviewee

Because the front end is currently in in my front company, the front end is quiet, quite tightly coupled with the AI stuff.

00:10:21 Interviewee

So, AI models and so they have to coordinate very often. So that's the basic. And then there is managerial duties were like I have team which I.

00:10:35 Interviewee

Leads. So, there are day-to-day activities like uh, which I have to manage without teams like one on ones with the each member of the team and then weekly joint meetings, then joint meetings and then some other.

00:10:52 Interviewee

Purpose oriented meetings that I have every week. So that's the basic role that I'm acting on right now.

00:10:59 Interviewer

Yeah. So, towards the question 5, do you have any experience in the previous company which is developing machine learning system? If so that what was your old experience?

00:11:10 Interviewee

No, in my previous company I had I haven't explored machine learning, but I did some. So I did create a project that was Rule engine. So, Rule engine is like something.

00:11:23 Interviewee

And step back from machine learning where you define the rules instead of data, but that was not really, I would say air stuff.

00:11:33 Interviewee

It was more like making the system capable enough to do multiple things at the same time, and it was so the.

00:11:43 Interviewee

Basic Edge and that system was.

00:11:46 Interviewee

Rules were defined by users and they can be written by users by like even after the programme was created and it was product like shipped to production.

00:11:57 Interviewee

So users can add and remove new rules, so that was the basic idea, and based on those rules the the.

00:12:06 Interviewee

System would do the inference on the data coming in data and give it like uh, classify the data has passed or failed, yeah.

00:12:17 Interviewer

So is your company service based or product based?

00:12:22 Interviewee

My current company service base, yeah.

00:12:28 Interviewee

The previous one was proposed.

00:12:36 Interviewer

So what software development model do you practise in your company? Like the?

00:12:42 Interviewee

In my current system, so our company I use agile in previous one we were working in with Waterfall where we would like have record documentation and then.

00:12:55 Interviewee

Like, I mean all three steps, documentation development, analysis, documentation and stuff.

00:13:02 Interviewer

So can you please share experience?

00:13:03 Interviewer

With the interesting projects that you worked on recently.

00:13:09 Interviewee

I would say with respect to the interesting thing was indeed the Starbucks project that I did because it was something that I did from scratch.

00:13:19 Interviewee

So we had some data about customers with the like we we the customers had.

00:13:28 Interviewee

Starbucks customer data, I I believe that that the data was not really the real data, but still it served the purpose.

00:13:35 Interviewee

So, the data will have customers and then there was these list of transactions that the customers did the transaction mean like buying or like buying coffee or getting some discount or anything like that so.

00:13:49 Interviewee

So, we had like 10 column UH-10 features or even less. I mean I think 8 features of customers and then even like five more features of each transaction.

00:14:02 Interviewee

So, we had to come up with a way to generate some data, like do some data engineering to generate enough features that.

00:14:09 Interviewee

Would help us create some good solution for modelling so.

00:14:15 Interviewee

So, there was a lot of work that we did on. I would call that I did on data engineering.

00:14:22 Interviewee

I had to, you know, link all the transactions in some way to generate some useful data. It was like at the end I was able to generate around 53.

00:14:34 Interviewee

Features for one customer and then on that basically on that that I trained KKK model sorry classification model that sorry cluster.

00:14:49 Interviewee

Model which clustered the data into like different categories and then on that clustering clusters I generated those clusters and then on that cluster clusters I actually did some heuristic to find out that what kind of like grouping is this. So is this group of customers.

00:15:11 Interviewee

Like more reason, more like more probable to buy if we offer them some discount. Or is it like is this group of customers won't buy at all. So this was this kind of clustering that we did and.

00:15:25 Interviewee

Then this. This was the basic outcome that we generated out of that. So yeah, that that was the basic design of the system.

00:15:34 Interviewer

OK. Thanks.

00:15:35 Interviewer

So, which common software architecture design techniques you find being used in most completely your experience in machine learning?

00:15:47 Interviewee

So the the architecture techniques that I have been using or we have been using in our company is mostly like we OOP is like object oriented design. That's the basic actually the technique that we used and after that there is test driven.

00:16:06 Interviewee

Development, though not test driven development because of the agile model, we do not really write tests before development, but we do try to add unit testing, add unit testing to our daily workflow and then the.

00:16:25 Interviewee

Then there are some very specific architectures like mono repo architecture and.

00:16:35 Interviewee

I think, yeah.

00:16:37 Interviewee

In in my previous company we would like use .net C#, so there are more design patterns that are used in there but in front end there are less design patterns because like for example React has its own framework and own designs to work with. So we do not.

00:16:54 Interviewee

Really have to do much of design.

00:16:56 Interviewee

That don't stop in that, yeah.

00:16:59 Interviewer

OK.

00:17:00 Interviewer

Yeah. So towards the the the 11th question according to your experience, what was your best software architecture design technique which are benefits of using them in machine?

00:17:15 Interviewee

Yeah. So, uh, I mean, uh, the best. I I up till now though though my I have my experience of functional programming is pretty like there has been a lot of experience in the functional programming in, in, in in light of react.

00:17:30 Interviewee

But not like on a back end technologies I have been doing front end functional programming front end so I have found myself more inclined toward a couple of things.

00:17:41 Interviewee

One is definitely typed systems. So instead of dynamic typing I think definitely type systems are more stable, more robust.

00:17:50 Interviewee

On the production production environments.

00:17:54 Interviewee

Like what C# provides us, and also what TypeScript provides us so like it puts out a lot of burden from our like testing to just compiler that does all the bug fixing that are very trivial bugs but they do get pushed to production when nobody cheques them.

00:18:13 Interviewee

So that's one then the second is object oriented design that I've seen applied not very often on front end.

00:18:21 Interviewee

But still it it has been used very often on back end technology. So object design does make it easy to visualise the system to for.

00:18:31 Interviewee

For someone who doesn't really know that intricate details of the system, so that's why with functional programming that gets a little difficult to.

00:18:41 Interviewee

Design like, make some diagrams out of that system to show to someone. But object oriented design does that like analysis stuff very easily.

00:18:51 Interviewee

So these are the couple of things that I think should be used like I will be using in future.

00:18:56 Interviewee

As well, until I find some thing better than that and then apart from that there are minor like practises that we do related to performance try to like try to make smaller functions and try to make smaller code.

00:19:17 Interviewee

Files than just large, like rather than writing a procedural code, write some object oriented code and stuff like.

00:19:26 Interviewer

That. Yeah, so yeah, nice.

00:19:31 Interviewer

Yeah. So towards the 12th question, do you have any recommendation for software architecture, design techniques of machine learning system?

00:19:39 Interviewee

So yeah, I have. I worked mostly on Jupyter notebook.

00:19:47 Interviewee

Which I found like interesting thing where I I use something similar for my local development rather local documentation. If I have to do I use.

00:19:57 Interviewee

Org mode in Emacs if. If you have an idea so it has similar functionality as Jupiter notebook where you can execute codes within the documentation, I mean you you you can execute the code snippets and get the result inside your documents. So that's the basic. So I do I I did like the.

00:20:17 Interviewee

That idea, but since Jupiter notebook is just a like kind of paper, you you have been given and you can write anything. So there's room for like writing something that is really not understandable so.

00:20:33 Interviewee

I I I haven't like found any alternative right now for Jupyter notebooks because since machine learning is a little different from our standard like development where we have to, you know, collaborate very often, machine learning does collaborate, but initial pass is always you know.

00:20:54 Interviewee

One one person doing everything or something like that, and then it's handed over to someone. So I think the industry is is still like in it's.

00:21:06 Interviewee

Like going toward maturity. And I think more tools would come in to.

00:21:11 Interviewee

Fill in the space that is there right now. I think the biggest problem that I see that can come up is like when you document something in a notebook, it's really not really easy to convert that to a programme right away.

00:21:27 Interviewee

Then you have to like convert and then if you convert that document into programme, a lot of information that in the documentation.

00:21:34 Interviewee

Uh, that is in the documentation. Get like is not translated down into the programmes, so I would say.

00:21:42 Interviewer

OK.

00:21:43 Interviewee

We need a better tool to, you know, do this Jupiter to code translation automatically, and then that would help a lot of a lot.

00:21:52 Interviewee

Of like machine learners, because machine learning engineer are mostly not more less interested in like production systems and how things are deployed.

00:22:04 Interviewee

And how things are running in the production, it's mostly the job of full stack engineers or someone who info like system. Sorry system engineers who managed the production.

00:22:16 Interviewee

So maybe this gap can be filled in the future, but I do not have any recommendation as such. If you call me right now.

00:22:26 Interviewee

Also, due to the fact that I haven't really worked on a production system of machine learning, I mean something production related. So might be my information is limited.

00:22:36 Interviewee

In that regards.

00:22:41 Interviewer

Yeah, fine. What would be the best practise that would could be useful or helpful in applying software architecture, designing of machining system.

00:22:53 Interviewee

Yeah, I think we can if we can like also find a way to implement object oriented design in machine learning systems. Some way to I mean I do know that Pytorch has this has this like class.

00:23:11 Interviewee

Called a base class where you can define a model with all the convolution.

00:23:17 Interviewee

Uh, but uh, like all the other systems like SLN and there is not much, uh, they they have other object oriented.

00:23:25 Interviewee

But when we do develop with them, we usually do not really care about objects. So maybe if we can more integrate more object into design in our training and training and.

00:23:38 Interviewee

Inference systems that might like help in future to make it more understandable for a general public, because right now the machine learning is something that.

00:23:50 Interviewee

Normal like software engineer would find little difficult to grasp on the first site like you get it. What is what's happening or what's not? Yeah.

00:24:02 Interviewee

Yep, Yep, that's all.

00:24:04 Interviewer

Yeah. OK. So towards the 14, what are the most common software design architecture design challenges in machine learning?

00:24:15 Interviewee

Uh for machine learning, the design challenge is always how do you? So it's the biggest architecture design challenge that I have like feel is there is how do you update a model when it's in production because most of more most I mean this is something like.

00:24:36 Interviewee

Since softwares the other softwares are not like that, they do not get stale overtime. But models do get stale over time with new data coming in.

00:24:45 Interviewee

So there's couple of things that is I think one is that I do not know about and the other is that I felt should work for the. So the first one is that I do not know about is.

00:24:56 Interviewee

How can we like update the model or like rather than just retraining it, just do something to put in new data and like?

00:25:07 Interviewee

Update within like previous model, updating on the new data instead of running all from the base, it just updates itself with the lesser computation power required.

00:25:18 Interviewee

So that's one I do not know about that is it already happening or not in the in the in the like machine learning equals.

00:25:27 Interviewee

Space. But that's one thing that I find. Like if that could happen, it would be very great in making maintainable software later on.

00:25:35 Interviewee

And then second is like even if we want to like retrain a model, how do we deploy that without like?

00:25:43 Interviewee

Getting the other model.

00:25:46 Interviewee

The turning off the other model. Yeah. Yeah. So, so integrating the new model in the prod.

00:25:52 Interviewee

While so the downtime managing the down time I have like we know that Sage Maker has some functionality around it, but like other than that I haven't like seen any tools or any techniques.

00:26:06 Interviewee

That would help us do that.

00:26:10 Interviewer

OK.

00:26:11 Interviewer

Yeah. What are the main architecture decisions on software architecture, design of different machine learning systems?

00:26:20 Interviewer

Like the the decisions really that you mostly find in these software architecture designs.

00:26:26 Interviewer

And or mostly in different machine image system like you have to do this then it will.

00:26:30 Interviewee

I think a lot of decision making in architectural sorry software machine learning is around data. So how many features you have to like have in this or in your training data and how many.

00:26:46 Interviewee

So as far as I can think, it's so it's mostly, I mean it's not really difficult to get those decisions done because in machine learning a lot of things are based on data itself.

00:27:00 Interviewee

So let's say if you want to know how many features you should go with, then you you do a PCA.

00:27:06 Interviewee

Of the your training space.

00:27:09 Interviewee

And then if you want to know which model you want to go with you, you actually benchmark multiple models along with each.

00:27:17 Interviewee

Each one of them like running the each model for 10% of your training data or 50% of your training data to reduce the training time and then you at the end just.

00:27:29 Interviewee

Improve your last like improve your most optimum model that you've received. So I think the decision making in our in software in and decision making in machine learning systems is mostly governed by data and.

00:27:43 Interviewee

If you want to think of architecture, it's right now there are like only few tools like since machine learning needs to be done on clouds, it's very difficult to do something like that on your own local system. You are you have very limited scope right now GCP is 1, then Sage Maker is the other.

00:28:04 Interviewee

And then there are these, like uh.

00:28:06 Interviewee

Other tools that we are using, for example, I don't know if hugging face offers this opportunity to you, but so I mean you understand like the tools we can use any which suits us with the pricing model.

00:28:20 Interviewee

Pricing is one of the reason you might go with Sage Maker or GCP or any other tool. So.

00:28:27 Interviewee

I think the mostly pricing and data is the governing part in the software design decision in machine learning.

00:28:35 Interviewer

Hmm hmm.

00:28:38 Interviewer

Yeah. So these all all the questions, all 15, so.

00:28:44 Interviewee

Yeah, I hope I'll, I'll did some useful.

00:28:48 Interviewee

Information to you.

00:28:48 Interviewer

Yeah, your your.

00:28:49 Interviewer

Feedback will be, yeah, your feedback will be useful when you try to combine all the practitioner.

00:28:56 Interviewee

So how how many? Uh, like people have you interviewed or are planning to interview for your thesis?

00:29:04 Interviewer

We are thinking to have at least 20.

00:29:07 Interviewee

OK.

00:29:08 Interviewer

Maybe I can stop recording them because.

00:29:13 Interviewee

Sure, sure. Yeah, that's fine.