Transcript

00:00:00 Interviewee

Like for sure.

00:00:02 Interviewer

OK, so yeah, I am recording after your permission so it has.

00:00:05 Interviewer

Started. Thank you.

00:00:07

No problem.

00:00:09 Interviewee

OK, you can like one by one. You can ask Me 2 questions. I am going through this and then I will be answering.

00:00:18 Interviewer

Yes, let me open the interview question here.

00:00:26 Interviewer

Yeah. So let me be introduced. OK, so can you please introduce yourself and describe your job role in your current company?

00:00:36 Interviewee

Well, I'm working as an software architect. Basically the role is.

00:00:43 Interviewee

Purely designing the softwares using the applied research at Inception Institute of Artificial Intelligence Inception, Institute of Artificial Intelligence is our national strategized organisation by "".

00:00:57 Interviewee

And they their aim is to basically develop a pure and applied research in the area of data science that can be transformed into real world problems, especially for the government of " " or for different kind of amortisation in within the country. So this is what the exactly.

00:01:17 Interviewee

Company aim is that is basically a national strategy organisation of artificial intelligence. So it's had, it is held by Professor Ling Shao as an advisory committee. We have Sir Michael Bradley.

00:01:33 Interviewee

Andrew and G as an Advisory Council member and and another is a person who is top ranked number one in terms of index factor in AI research and computer vision. Who is Sir Anil K Jain is a professor full professor in MIT.

00:01:53 Interviewer

Oh wow.

00:01:53 Interviewee

So these are the Members who are basically the part of Advisory committee. In addition to this, the inception Institute of Artificial Intelligence.

00:02:04 Interviewee

Working with first specialised University of Mohammed bin Zayed artificial intelligence at Abu Dhabi, which is the first institution in the world who is only providing the like capability of MSNP PhD in AI. So we are we are basically.

00:02:24 Interviewee

The Institute and Mohammed bin Zayed on the same floor and we have joint ventures, different kind of projects, different kind of even our faculty is shared among institute research as well As for the academia. So this is this is basically the Institute.

00:02:43 Interviewee

As I like told earlier, the Institute basically provides the background and the back end play research as in subsidiary company of G42G42 is one of the company that is AI company in.

00:03:04 Interviewee

As a royal part of Royal Group, Royal Group is owned by the Royal family.

00:03:09 Interviewee

So this, that's why this initiative was like taken by the Royal Group and in Back 2018. So my role is basically.

00:03:23 Interviewee

To work as a bridge between.

00:03:26 Interviewee

Researchers and the business development consultant people, as well as the engineering team.

00:03:33 Interviewee

So I am the only one in inception who is data science and machine learning architect who is responsible to reshape and transform.

00:03:44 Interviewee

All apply research into real world problems. For example, we are doing research in different kind of medical units.

00:03:56 Interviewee

Health units.

00:03:58 Interviewee

As well as in satellite imagery.

00:04:02 Interviewee

Spread imagery and monitoring and controlling for the government of Abu Dhabi Police Department, specially Municipality and transportation. So these researchers mainly work as a background applied.

00:04:18 Interviewee

So what I do is mainly I work and do some meetings with different business development people, with clients, with government units and based on these meetings and M's I find or expect some data that can help me to providing some solutions.

00:04:39 Interviewee

To provide some solutions to what the real world problems.

00:04:44 Interviewee

So this is this is 1 responsibility. The second responsibility is basically.

00:04:49 Interviewee

I need to know the technology stack. I need to know the machine learning area and I need to design all these kind of stuff for real world problems and their solutions.

00:05:02 Interviewee

Currently I am doing I did last year 9 assignments in different areas, mainly in transportation monitoring and controlling and municipality.

00:05:11 Interviewee

Nowadays I am working on Dubai Expo which is a very big assignment and we have automation of detection of vehicle all over the Expo.

00:05:21 Interviewee

Right. And there are approximately spots are more than 50,000. There are fields which are we are taking from more than 5000 cameras.

00:05:31 Interviewee

And based on that, we are applying designing architecture technology stack and each and everything end to end solution under like software design patterns under software design methodologies and we are developing using the AI innovative technologies and cutting edge technologies. So this is what my role is. What I am involving.

00:05:51 Interviewee

I'm being involved in different kind of things.

00:05:53 Interviewer

Yeah. So right now you are working on the image processing for vehicle detection.

00:06:00 Interviewee

Image object Detection, object tracking, and and segmentation and classification as well. There's a deep learning based machine learning, deep learning and reinforcement learning, meta learning as well. But mainly there are two areas which is classical machine learning and image processing plus deep learning.

00:06:04 Interviewer

OK.

00:06:22 Interviewee

And it's it's I think it's computer vision.

00:06:22 Interviewer

Are you using any?

00:06:23 Interviewer

Cafe. Are you using any cafe or something like for deep?

00:06:27 Interviewee

Yes, yes, yes. Yes. Uh. For Edge computing models, I use Caffe, Tensorflow, Pytorch and the state-of-the-art model we usually pick from our own in-house researches as well as from Facebook sense.

00:06:43 Interviewee

And Baidu and the Berkeley BDD 100K data set, which is a purely data designed for vehicle detections. So so we use the latest analogies.

00:06:55 Interviewee

In in like in terms of programming, in terms of research, in terms of machine learning, the benefit is because I'm the person who is who knows the pro and con.

00:07:06 Interviewee

Or knowledge of software architecture and design and software engineering as well as machine learning. But the people who are purely belong to the research area, they provide me a very good feedback in terms of latest models and latest papers, latest research, latest technologies in purely in deep learning. So that helps me a lot in terms of.

00:07:26 Interviewee

Designing and making some systems for different entities.

00:07:31 Interviewer

Hmm yeah.

00:07:33 Interviewer

So how many years have you working right?

00:07:37 Interviewer

Now in this company.

00:07:39 Interviewee

Is approximately going to be 3 three years. I landed in Abu Dhabi on June 2018. Since that I'm working in the same domain, though previously there was one company ALGORITMA, but that was.

00:07:53 Interviewee

Like as a again that was owned by a DFG which is again the Royal Group.

00:08:00 Interviewee

But but we can say that I I always say that I'm working with inception since last three years.

00:08:06 Interviewer

OK, so how many total experience do we have?

00:08:10 Interviewee

My my professional career started in 2005 March, but in big data I moved in 2015.

00:08:20 Interviewee

To till 2018, 14 to 2018, I was in big data and Pure data sets and machine learning area. I started my career from June 2018.

00:08:41 Interviewer

Yeah. So have you published any thesis research?

00:08:44 Interviewer

Papers. Machine learning.

00:08:47 Interviewee

No, in machine learning area, to be honest, I don't have enough time to like publish the papers. And because my mainly focus is to deliver some tangible or like the platform, the solutions. So I myself did not delivered any publications. However my master.

00:09:07 Interviewee

Pieces or I had a paper in in 2012 that was.

00:09:13 Interviewee

There was basically distributed systems and using formalisation or formal methods, but that is only one publication in Pakistan and the thesis is obviously in in MSC and Ms in both.

00:09:26 Interviewee

I I did my own thesis.

00:09:28

OK.

00:09:29 Interviewer

So do you have any experience in the previous company which is developing?

00:09:32 Interviewer

Machine learning system. So what was your old experience?

00:09:37 Interviewee

No, it was it. They were all classical methods of AI. We cannot say they are purely machine learning because the pure deep learning started in 2012 around that. But mainly they work on classical methods which are purely mathematical methods, computer vision methods.

00:09:56 Interviewee

But since 2000, 2018, my this is my first job in pure machine learning or deep deep learning areas.

00:10:03 Interviewer

OK. So is your company service based?

00:10:06 Interviewer

Or product based.

00:10:08 Interviewee

The both in certain area we provide services and in certain area we provide our platforms or like like we have this is 1 solution that I am working at the moment is a smart parking solution, smart monitoring solution and smart tracking solutions for the vehicles and persons. So this solution is.

00:10:30 Interviewee

Purely IPO of IA. In addition to that, we have our satellite imagery platform our own, but that is a collaboration with the Huawei as well as the GEOENGINEERS and L3, Harris and ISRI as well. But in certain area we are providing just consultancies to the different departments.

00:10:50 Interviewee

Of " " like we, we have different joint venture in the health department. We have different services with Adnoc Abu Dhabi or Gas Corporation. So in certain areas consultancy in certain area we have our own platforms.

00:11:05 Interviewer

Hmm. OK fine.

00:11:07 Interviewer

So which software development model do you practise in your company? In general, like Agile or Waterfall?

00:11:15 Interviewee

It's again, again, mainly because it's more awareness towards agile and scrum. So mainly units where I work with the different departments and that we can different teams. So there we we mainly focus on agile and.

00:11:34 Interviewee

But personally, uh, the projects I do, I don't follow the programme. I mainly work on the spiral model.

00:11:43 Interviewee

So in certain areas, the waterfall as well because we are POC's and we are providing the demos, we are definitely we we do some prototyping model as well.

00:11:53 Interviewee

But where we have a big project and we don't have bioco demos there, we are following Scrum and Agile in healthcare in.

00:12:03 Interviewee

Defence systems as well as in smart nations platforms, but in certain areas where we have some research involved where we are different kind of methodologies are involved. There we follow spiral in inception.

00:12:18 Interviewee

Server based on ad hoc base or based on need. We usually use waterfall and prototyping model where we have to deliver some POC or demos.

00:12:29 Interviewee

As well.

00:12:31 Interviewee

So it depends upon different because it's a totally hybrid hybrid environment. It's not like we are tightly couple where the teams are more organised, well established.

00:12:42 Interviewee

We have a clear requirements there. We usually follow a jar and scrum where we don't have. Then we go with the waterfall and prototyping always.

00:12:49

OK.

00:12:51 Interviewee

Oh yes, spiral spiral is again is ugly. Spiral is also for the mature and stable area, but it's old classical method and software engineering, especially the software development life cycle where we have involved scientific projects, but mainly it's at the moment and agile is all over the industry is.

00:13:07 Interviewer

OK.

00:13:11 Interviewee

Famous at.

00:13:12 Interviewee

The moment.

00:13:12 Interviewer

OK.

00:13:13 Interviewer

So could you please share your experience with interesting projects that you worked on recently? I think you have already explained.

00:13:19 Interviewee

Yes, yes, that that's there are three to four players that I work. One is we have a mega prayer with the integration with the Hava Geo Genius, which is applied imagery. I provided the designer architecture of distributed jobs just like MapReduce and Kafka.

00:13:34 Interviewee

And we designed this basically end to end pipeline with the with the Geo Genius platform. And the second is basically the monitoring and controlling for municipalities and smart nations.

00:13:46 Interviewee

And the third one that is is I'm currently working is the Dubai Expo provide which is a mega provide and is a big assignment where we have a lot of cameras.

00:13:55 Interviewee

Vehicle classification, segregation, segmentation, even we are applying there.

00:14:00 Interviewee

Grappy projective transformations blah blah like that. Even the GIS system is also involved in that because it's a national level, even not national level, because the international level event.

00:14:12 Interviewee

So that's why it's important and level of critically like criticality level is too high. So that's that's why this is really.

00:14:21 Interviewee

An interesting project as well.

00:14:25 Interviewer

Is your working experience? How many software architecture, design techniques you're working?

00:14:31 Interviewer

In machine learning.

00:14:31 Interviewee

Overall machine learning, it's not more than three, three years. It's it's fully with the three years dedicated because I I'm working since last June 2018.

00:14:36 Interviewer

OK.

00:14:38 Interviewer

OK.

00:14:42 Interviewee

So it's it's my first job was at a software architect. So as a data science architect or machine learning architect so since 2018.

00:14:50 Interviewee

Like I'm working in machine learning solutions, the selection of technology selection of tools. For example I I give you an an idea use case. Suppose while doing some preprocess.

00:15:04 Interviewee

Thing where we have ample of time, we have ample of computational power, then we usually do in math in Python or some or or some like such kind of languages where doesn't matter performance or architecture it like we can't afford latency. We can afford some bugs. We can afford some like.

00:15:25 Interviewee

Pipeline breakage that that can be affordable, but when we we deliver that solution or model into production then I we mainly focus.

00:15:33 Interviewee

On C plus.

00:15:35 Interviewee

And mainly efficient methodologies like designing what kind of database it will be involved, the back end side, what kind of inference server will be involved.

00:15:44 Interviewee

That is the main fundamental thing in terms of software actually design that we need. You really need to focus and we will need to think about that how my solution will work and how and how.

00:15:55 Interviewee

These systems will integrate and communicate with each other.

00:15:59 Interviewee

And the main atomicity and cohesion and coupling that is a fundamental 3 parameter, they are we need to keep in our mind all the time, whether it is machine learning, whether it is not machine learning, but mainly in, because this is the whole area is open source and where the open source areas are involved we we have.

00:16:19 Interviewee

Like lot of issues while integrating the systems because the every person might be.

00:16:26 Interviewee

Person A is developing their solution or model which is related to detection that is only compatible with the Pytorch.

00:16:35 Interviewee

And the other one is the for the classification might be might be that will be only compatible with the TRT or tensor or tensor flow.

00:16:42 Interviewee

So you have to think about that. So what kind of technology and how much accurate that model is and that will be useful for me and what kind of technologies I I need to implement or suggest?

00:16:55 Interviewee

For designing or implementing that solution, so there are many mainly important. This is 1 area that is the pre like pre deployment area and the second area is like deployment area where we need to think about.

00:17:08 Interviewee

The scalability performance and its extendability of the system that is main fundamental thing and I believe at the moment we have a lot of research, we have a lot of things that that are improved by the researchers. But unfortunately this area, the scalability, extendability, integrity, integrity of the.

00:17:28 Interviewee

Data and and it's the enhancement of the system where the AI area is lacking.

00:17:34 Interviewee

A lot.

00:17:35 Interviewee

This area needs to be improved and the bad thing is, unfortunately, it's my own experience. Maybe it's not all over the world, but it's my experience, our mainly the people are decision making. People are involved who have good research background.

00:17:52 Interviewee

And they think everything just like this is a part of research. But this is not the case. Designing the solution designing the software is a totally different thing.

00:18:00

Thank you.

00:18:01 Interviewee

And providing a model of aibs model that is a different thing though all things should be aligned with each other, but their mentality and mind of thinking even I am working with.

00:18:13 Interviewee

The researchers they.

00:18:14 Interviewee

Are really competent, very capable, but unfortunately, when we are going to design a solution or system.

00:18:23 Interviewee

At Mega or at Mega skill level?

00:18:26 Interviewee

Then they don't know what kind of stack I should use and what kind of inferencing platform I should use, what language I I should be I should select for, for, for for that where it should be best. So from that area, based on their mindset.

00:18:44 Interviewee

Because they are the leading at the moment. It's it's I I will call it. I it's it's my personal opinion. Maybe some people don't agree. It's like out of this.

00:18:52 Interviewee

The talk is out of this box. Some people think that because this is a phobia of that, people who are leading, who are doing PhD and they they have very competency in research, but they are leading the systems. But this is a bad approach.

00:19:10 Interviewer

Yeah, I think so too.

00:19:11 Interviewee

Solutions and software always led and they should be like designed or their their level of thinking should be the people who are really capable in in software engineering and they must be aware of design principles or software development principles.

00:19:30 Interviewee

So here the researchers have deficiency.

00:19:35 Interviewee

They are not capable to do that job and this is to be honest, this is not their job. So this area is really lacking and where it's my own opinion, I I advise my researchers and my boss and my the colleagues are where I am involved. I always suggest them to.

00:19:55 Interviewee

Give the right job to the right person so that they or he can think or full stack developer can think better than the person who is the back end developer or front end developer.

00:20:08 Interviewee

So definitely a person who is the expert in the node, JS or JS. They he can do the better job than the person who is expert in C.

00:20:15 Interviewee

Person I am working in since like last 10 to 15 years in the C language and now I'm last few years I'm involved in the Python so I know pro and cons of the both where I should use where I should not use.

00:20:27 Interviewee

So this is what kind of things are. So this was like out-of-the-box discussion that I was doing.

00:20:33 Interviewer

Yeah. So, yeah, I I would also agree with you like we we might need the research and the developers perspective didn't have collaboration and they came up to the same solutions before.

00:20:43 Interviewee

The bad thing is the bad thing that I observed in the market is in the AI area. The people who doesn't have experience in software engineering.

00:20:53 Interviewee

But they never designed the solutions.

00:20:56 Interviewee

Once they do their PhD, they came into the market and they perform leading leading job leading job in a.

00:21:04 Interviewee

They can do the leading job in research, whether they have very good publication or CPR, IEEE and other.

00:21:12 Interviewee

Known conferences and and like the journalist. But The thing is.

00:21:20 Interviewee

They are good in research, but this doesn't mean that they can design a very good software solution. So there it's it's. It's like like in my company all leading positions.

00:21:32 Interviewee

They are led by people who are PhD in computer Vision AI.

00:21:37 Interviewee

But they have.

00:21:40 Interviewee

Three to four years industry experience.

00:21:43 Interviewee

Where a person who has a 15 years of instruction, he knows much better than them. A software design, even the the management should think, OK if one person has delivered a continuous track of delivery of the software, he knows much better than the person who.

00:22:00 Interviewee

Recently graduated from a known university, may Oxford, MIT, Stanford, UC. I don't mind that I'm. I'm like, good. I learned from them a lot.

00:22:11 Interviewee

But the bad thing that I feel is from software design, engineering and software methodology perspective that is a bad thing. They these kind of things and the decisions.

00:22:20 Interviewee

Are led by the people who are really competent in research, but they are doing their decisions in software engineering, which is a bad.

00:22:34 Interviewer

Towards the dance question, according to your experience, what are the best software architecture, design techniques and what are the benefits of using them?

00:22:41 Interviewer

For machine learning.

00:22:46 Interviewee

From I will talk on generically like software architecture and design methodology that mainly are useful. One is the client server architecture.

00:22:58 Interviewee

That is a fundamentally is a very good and it is being used and it should be used #1 #2 whatever the model we use, whatever the architecture we use, we must keep in our mind.

00:23:12 Interviewee

Coupling cohesion. That is the fundamental point.

00:23:16 Interviewee

And like if if we we are designing, we must be aware that inference in in the in the machine learning area, the inference part should be totally isolated and generalised.

00:23:28 Interviewee

#1 #2 data.

00:23:32 Interviewee

Better part that is.

00:23:35 Interviewee

Feed to the server or as an output of the server that must be isolated and behind a firewall.

00:23:42 Interviewee

For the data integrity and and that kind of stuff, and there should be some stable isolated skips or algorithms, they must produce data. So this is a generic architecture that we can say it's a client server.

00:24:00 Interviewee

Architecture that mainly I because my experience I have been working in since long time in the client server architecture.

00:24:08 Interviewee

So I mainly prefer in that with the low cohesion. Sorry with the high cohesion and low coupling, high cohesion and low coupling and data integrity.

00:24:19 Interviewee

Integrity part and data security as well.

00:24:22 Interviewer

OK.

00:24:23 Interviewee

Now the next next is in the in in the theme is that I feel is it should be that we must learn some.

00:24:32 Interviewee

Existing architecture and design with respect to the use case that we are going to implement or that we are going to achieve in that ML solution.

00:24:42 Interviewee

Whatever that architecture is, it doesn't like. It's very flexible. Whenever there are open source technologies, we are flexible to achieve any of the goal. So it depends upon whatever solution you are going to build suppose.

00:24:56 Interviewee

We have time constant.

00:24:59 Interviewee

Then we cannot go with a system or architecture that have.

00:25:05 Interviewer

All time OK.

00:25:05 Interviewee

That we can afford that latency, we need to work in the real time mode like we cannot afford the post and get request. We have to establish a TCP IP or GRPC string.

00:25:16 Interviewee

And similarly the solution or the system have different kind of modular approach that modular architecture should be involved in that and every module should be responsible to their own.

00:25:29 Interviewee

So these are the mainly in generic terms that we use in like in a software architecture design. And I always suggest and I always work on these, obviously there are fundamental principles or paradigms underlying that software engineering that we use. So they are like some kind of theoretical things.

00:25:49 Interviewee

That we always implement and try to implement that contains data entity atomicity, fall back redundancy and client server architecture and different kind of paradigms and design patterns as well they are all.

00:26:03 Interviewee

Under any of the architecture we can implement, but depending upon where we need the best, if we we can afford the time then the architecture we need to follow is not client server.

00:26:16 Interviewee

We can afford as a standalone model with the high and low coupling and if we have not enough time to.

00:26:23 Interviewee

Maintain the data, then obviously preprocessing you do the best in the initial development level, and then we use the client server architecture or TCP team or some any of the other communication channels for any kind of solution.

00:26:39 Interviewer

Yeah. OK.

00:26:41 Interviewer

Yeah. So do you have any recommendation for software or texture design techniques of machine machine learning?

00:26:49 Interviewee

Yes, I have some advised and that I you will provide to my own engineering team as well and the research team as well. So fundamentally.

00:27:02 Interviewee

Fundamentally, based on the requirement.

00:27:06 Interviewee

We must.

00:27:08 Interviewee

Have some comparative analysis.

00:27:12 Interviewee

Of software architecture and design, while while picking one of the architectures #1.

00:27:20 Interviewee

So we must have some comparative analysis and then then the next is we use weather data-driven approach or test driven approach.

00:27:30 Interviewee

We must have some.

00:27:32 Interviewee

Data at the initial level.

00:27:35 Interviewee

Or we must have some requirements to make a decision?

00:27:40 Interviewee

So this is the fundamental that I use, but mainly when we do.

00:27:47 Interviewee

Any of the solution? Are we going to implement any of the solution that I always suggest to use client server architecture.

00:27:56 Interviewee

And atomicity in every module so that we should not be like will not be going to break any pipeline or any interruption with other modules.

00:28:22 Interviewer

OK, So what would be the best practise that could be useful or helpful in appliance software texture designing for machine learning?

00:28:34 Interviewee

If we are going to.

00:28:37 Interviewee

Do with the push your demo then obviously I will suggest the prototyping and waterfall model. Follow that practise.

00:28:45 Interviewee

With the continuous integration or with the continuous interaction with the.

00:28:49 Interviewee

Client and if we have complete requirements and then obviously we should go with the agile and come with the continued integration of the countries interaction of the client.

00:28:59 Interviewee

So to ensure that everything is going on the right way, so this is this would be the best practise.

00:29:05 Interviewee

I think for the machine learning areas to.

00:29:08 Interviewee

For good software attraction design.

00:29:10 Interviewer

Yeah. So you think so that it would be also nice to have a component architecture or required architectures?

00:29:17 Interviewee

Yes, obviously, obviously, obviously see software engineering is not like if what I'm saying is is hard and fast. Obviously what I'm saying it's my based on my own experience might be there a lot of other areas like NLP.

00:29:32 Interviewee

Which is there is a graph knowledge as well as there are time series problems, so these are the other areas.

00:29:40 Interviewee

So maybe I am focused on only computer vision detection and classification segmentation area. So based on based on that area, maybe my my perception will be different before that if I say pure software engineering method.

00:29:45 Interviewer

OK, OK.

00:29:52 Interviewee

Obviously there are different kind of we can say.

00:29:56 Interviewee

That still we did not discuss any QA in this, the whole that is a fundamental part of that we did not discuss the requirement engineering part which is the pure part of the software engineering design patterns and architecture and.

00:30:10 Interviewee

Design. So these are all components are involved based on like if you are taking interview for my opinion.

00:30:16 Interviewee

From this side I will be giving you five points, 5 basic points that that will be useful for your data. The next rest discussion will be maybe.

00:30:25 Interviewee

Similarly, you will be going to interview a person who is more experienced or more different. He is working in the different angle then he will be providing three or five more additional points and when you will merge all these points, you will be concluding OK, these five points discussed by Ajmal and these discussed one by they. If we merge them.

00:30:45 Interviewee

We have a concrete software engineering architectural design. So what my area is I am working on that area.

00:30:53 Interviewee

Again, it depends upon if your company is a consulting company, then the paradigm, the methodology, the processes, the shops are quite different, and if the company is the product based company then it's totally different. If I go in 2006 and seven when I was a part of Halliburton LKR.

00:31:13 Interviewee

There we usually use a spiral model and there was no agile, no scrum. It was not in birth at that.

00:31:21 Interviewee

That time, so only we were using the conventional waterfall model or parallel model. Mainly even we have a very good team of the and we usually main stop like software services oriented architecture at that time.

00:31:38 Interviewee

So micro microservices as well. So microservices still work at the moment.

00:31:43 Interviewee

But depending upon.

00:31:45 Interviewee

In my area where I am working might be it's not best because I'm mainly focusing on client server architecture, maybe services oriented architecture is also the client architecture is again a micro kind of micro services architecture. Yes yes. But my focus is or my thinking is.

00:31:59 Interviewer

Language processing and other thing.

00:32:06 Interviewee

If we are involving huge number of dockers or huge number of services that it will reduce your performance and overall system capability.

00:32:16 Interviewee

So always we need to think while designing a solution in the ML it's a very good approach to use microservices architecture services.

00:32:25 Interviewee

Oriented architecture or client server architecture, but you have to make decision according to the problem statement.

00:32:32 Interviewer

Profiles and requirements, yeah.

00:32:33 Interviewee

Yes, problem requirement. They're a fundamental thing.

00:32:37 Interviewee

And a person who is responsible to design provider, designer architecture. That's why it is always recommended he must be capable of delivering in different diversity in the the different dynamics.

00:32:49 Interviewee

If I am working on only one domain, I'm not a.

00:32:52 Interviewee

Good software architecture.

00:32:54 Interviewee

I'm not a good software, I'm not.

00:32:55 Interviewee

A good designer.

00:32:57 Interviewee

I could be a very good software engineer or a good good programmer, but if I am working since last 15.

00:33:02 Interviewee

Three years in one domain, then. Then I never have a taste for the Java technologies. I never had a taste for the spark streaming, but if I am using Spark streaming Kafka and different kind of then I must have some knowledge at least at that level. So I can decide yes last time.

00:33:19 Interviewer

And what is the different vantages?

00:33:22 Interviewee

One person he is a leader, head of data science for our healthcare department. He came to me. He was saying that we are going to hire a person who is the technical architect.

00:33:33 Interviewee

So what should I do? I said OK, hire a person like who is he? They they gave me a seven. I introduced him. He was from delight India. But he this person was really good.

00:33:43 Interviewee

He was expert in AWS. He was very expert in Python but.

00:33:48 Interviewee

From that circle, he has no knowledge. He has nothing knowledge in the software design principles. Even he only have one or two approaches that he he had been using since last 15 years, because his

tenure in delight is more than 10 years. So we cannot say that if he is a leading leader for that company, he is.

00:34:08 Interviewee

Poor man, but for another company, maybe for my G42 or my XYZ company, you would not be a good person, so it's always recommended to have the best knowledge, breadth wise knowledge, deep knowledge definitely matters in one area, but it is not mandatory to have deep knowledge.

00:34:28 Interviewee

In all area.

00:34:29 Interviewee

Breadth knowledge must be there. Like tools, technology selection and each and everything there must be there for a person.

00:34:36 Interviewee

And the one thing that I thought I I don't want to miss is that this kind of thing is really lacking in AI at the moment.

00:34:46 Interviewee

It's it's. It's like what to does it, to be honest.

00:34:52 Interviewee

I always criticise in my own company. This area really needs improvements and if we AI has done a lot research is like all over the world is going drastically in AI.

00:35:07 Interviewee

But unfortunately, 99% research is going to be a rubbish or parish or we can say that it is going to the dustbin because only 1% solution there Google or Facebook, they are really transforming into our right solution into the right area where we can say, OK, I gave you an example of we have a detection system.

00:35:27 Interviewee

Detection system where I can use I can use in many areas but.

00:35:32 Interviewee

It's used as daily. We are getting more and more and more publications but.

00:35:37 Interviewee

Improvements are there in the detections, but no one can say that this paper is really transformed into a well scaled system, and this system is going to be deployed in XYZ location.

00:35:50 Interviewee

In my own project for the Dubai Expo.

00:35:54 Interviewee

The AI part is only 10%.

00:35:57 Interviewee

The AI part is only because what it is, what it is doing at.

00:36:01 Interviewee

The moment is, it's just.

00:36:03 Interviewee

Giving the image to our server and getting the detections back. That's it.

00:36:08 Interviewer

That's it.

00:36:08 Interviewee

Rest of the all data is 9990% more than 90% is the work that I am doing in in the front end the back end, the database, the query schema.

00:36:20 Interviewee

GIS homography the projection transformation that long blah blah like they're all these areas and these components must they are all belong to pure software engineering methodologies and if they are well designed in terms of database the again the other area that we forgot is the database database is a very broad area but we never like.

00:36:40 Interviewee

Provide a lift to the people who are doing the database engineer so that is again a database architect who is like purely a very big area and this is the basic domain in in the computer science, which is a purely computer science domain.

00:36:54 Interviewee

Rest of all domains doesn't belong to the purely computer science domain database is the purely domain which belongs to the computer.

00:37:01 Interviewee

Rest of are integrated with one of the other domains like if I say.

00:37:05 Interviewee

AI is a merger of computer science. Plus mathematics networks is electrical engineering plus computer science. Similarly, we can say that computer graphics, mathematics and computer science.

00:37:15 Interviewee

But database is only the computer science domain, but that we never think like that on that area. There will lack of the people doing.

00:37:25 Interviewee

Doing solution and the pickle dumping Jason format but that is when you are going to deploy the solution in a big level the with the respect to microservices or micro service or whatever we follow.

00:37:38 Interviewee

The system doesn't work for the huge number of the million for.

00:37:40 Interviewee

The 1,000,000 transactions.

00:37:41 Interviewer

Yes, exactly. Yeah, yeah.

00:37:42 Interviewee

So that area is basically need to be improved and apart from the research and from your topic.

00:37:51 Interviewee

It's it's kind of vision, lack of vision, knowledge and thinking of the higher management who are really working in AI area.

00:37:59 Interviewee

For example, my company CEO never hires a very good software manager. He always tried to hire a very good researcher.

00:38:07 Interviewee

So this is a wrong approach if you are going to build a company with the product on one side, he says.

00:38:13 Interviewee

I need an excellent product. I need an excellent solution on the other side, he hired a person for that solution. An excellent researcher.

00:38:23 Interviewee

Never an accident software manager, so this kind of thing is there. That's why, because this is my professional into with you for your research.

00:38:32 Interviewer

Yeah, yeah.

00:38:32 Interviewee

So I definitely want to convey my message. What I am feeling as, but definitely you can in take an input from the others as well what they think, I don't know, but this is what my area or my expertise that I am.

00:38:34 Interviewer

Yes, exactly. Yeah, sure.

00:38:43 Interviewee

Going to blow.

00:38:44 Interviewee

You is. That is my birthday.

00:38:44 Interviewer

Yeah, yeah, yeah, your your.

00:38:46 Interviewer

Input is important.

00:38:47 Interviewer

To me, yeah.

00:38:48 Interviewer

Yeah, towards the 14, what are the most common software architecture, design, machine learning system?

00:38:55 Interviewee

Most common software architecture design challenges in the machine learning one major challenge is the scalability.

00:39:02 Interviewer

OK.

00:39:03 Interviewee

#2 is integrity.

00:39:06 Interviewee

#3 is obviously the cohesion and coupling is the fundamental thing and and and the other is like they don't follow purely microservices or software and architectural services in that at the moment they they are at this time.

00:39:23 Interviewee

And some apart from that, there are many others. They are the integration continuous integration problem. Is there is a very big thing in the machine learning area has continued integration. Certain one there is no flexibility in the existing architecture that can support or afford or integrate.

00:39:44 Interviewee

A model of the other person.

00:39:47 Interviewee

So there is a lack of hell of lack.

00:39:51 Interviewee

So these are the major challenges and the other is challenges like data availability.

00:39:58 Interviewee

That that part, yes, that availability and the labelling of data if we talk about more related to the research area though, this doesn't belong to software engineering, software engineering is nothing to do with that. But just as a topic, I would like to discuss from machine learning area, the challenges are the.

00:39:58 Interviewer

Yeah, continuous your data availability.

00:40:18 Interviewee

Availability of data.

00:40:20 Interviewee

Labelling of data annotation of data. So these are the three labelling or annotate annotation of the data.

00:40:25 Interviewer

Annotation OK.

00:40:26 Interviewee

Annotation annotation labelling is the same thing.

00:40:30 Interviewee

First, the labelling and annotation. Is there any person you or me or any person he can train the model because we have a lot of architectures are involved in there. So again depending upon the problem statement. So these kind of challenges are existing in there and.

00:40:46 Interviewee

One major challenge is again, I would like to say is a vision.

00:40:50 Interviewee

And breadth wise knowledge.

00:40:53 Interviewee

The people have a very deep knowledge, but they don't have a breath knowledge. If I want if I say OK this model cannot predict a particular person.

00:41:06 Interviewee

Then engineer or software engineer can.

00:41:10 Interviewee

Think out-of-the-box. But Sachin will never think out-of-the-box.

00:41:16 Interviewee

Researcher will never do such kind of reverse engineering to resolve the problem. So there there are many hacks or many points where we as a software engineer we need to resolve it by using out-of-the-box approach.

00:41:33 Interviewee

So that that is another challenge if we if the model is not giving you right out.

00:41:39 Interviewee

So what you will give it to the client, you will say that you can never say to the client that my model is not working because he has nothing to do with the model.

00:41:48 Interviewee

You always need a solution. My client is the Expo. He is saying like, OK, your model should work to detect right position. That right place for the vehicle or the person. So I cannot say my my my model.

00:42:00 Interviewee

Is not giving me the right detection of the right segmentation. The right output. He has nothing to do with that, so I have to do some paradigm or some approaches that can resolve the client problem.

00:42:11 Interviewee

So this is another challenge and as I elaborated, the vision problem is there data problem is there and cohesion and coupling problem is there scalable scalability and enhance enhance that is the problem and and and other problem that is related to deployment.

00:42:30 Interviewee

We have a lot of frameworks involved in the deployment.

00:42:35 Interviewee

We cannot say that that they, they are as exactly fitted to.

00:42:41 Interviewee

To a solution because if I'm using the Kubernetes for one deployment, might be that communities is hell of expensive in terms of time resources. Whether why should I use the doc simple Docker?

00:42:55 Interviewee

This and and other challenges like if I'm going to use.

00:43:00 Interviewee

Use a big Apache web server. Why not I should use if my there are very minimal number of requests I have to come on platform that is.

00:43:07 Interviewee

A flask I should use an open API or a flask kind of thing?

00:43:13 Interviewee

So these kind of challenges are also there to in, in terms of selecting our technology and tools for deployment as well.

00:43:20 Interviewee

And the other is basically this is called ML OPS. So in generically drops ML OPS is just like a dev OPS.

00:43:29 Interviewee

Just like we call it Dev OPS, it's now that term is introduced as an ML OPS machine learning operations kind of job and support.

00:43:37 Interviewee

This area is really lacking at the moment. No one is trying to jump into that area.

00:43:44 Interviewee

For a solution.

00:43:46 Interviewee

Continuous integration, continuous support, dev OPS, ML OPS. These areas are evolving at the moment.

00:43:55 Interviewee

So these are the general challenges there.

00:43:56 Interviewer

Yeah, yeah, very interesting.

00:43:59 Interviewer

Yeah. So towards our last question, what are the main architectural decisions of software architecture design?

00:44:06 Interviewer

Of different machine learning system.

00:44:10 Interviewee

I think it's the same that we we discussed. You can conclude this answer that as I while giving you the statements, I was also giving you the advice.

00:44:21 Interviewer

Exactly advises.

00:44:22 Interviewee

Based on that, you can conclude that that if we are going to design a system, So what the best architecture we can follow and what the decisions we can make fundamentally it's it's like like we when we have.

00:44:39 Interviewee

Degree we do some graduation from some we have a basic knowledge similarly that basic knowledge is implemented everywhere in every computer science domain. So software engineering from software engineering perspective.

00:44:51 Interviewee

The I can say that it's a bitter truth and it's unfortunate.

00:44:58 Interviewee

That the people who are really graduated from higher university in, in the area of computer vision AI, they never think or they never have a good.

00:45:07 Interviewee

Giving a good space to the software engineering people and they are not giving a good space to the people.

00:45:13 Interviewee

Who? Who talk about the software methodology, even even sometimes myself, I'm opposed to them. That, oh, no, no need. We just need to focus on on.

00:45:25 Interviewee

Technical area that I will miss one point as well. The documentation area and the cube area the documentation, the queue area because we are discussing and I am also the back end my my mind is giving me thread and info that which areas are we are lacking at the moment.

00:45:41 Interviewee

So we need to have some decisions, not only from the requirement engineering, not only from the design methodologies, not only from the like comparative analysis of the architecture for the for the particular problem statement.

00:45:54 Interviewee

But also we must have some well documented applications as well.

00:46:02 Interviewee

I still found I still find many publications, many papers, they are very good, well documented. I don't want to talk about that because our researcher can understand better than me.

00:46:13 Interviewee

But from software perspective, if I find a repository or GitHub code, I cannot run that code. That is another challenge that maybe I would like to expose as in as in client of a researcher I am a client of a researcher because I I use his research to transform into an application.

00:46:33 Interviewee

So what they usually do.

00:46:35 Interviewee

There's the the either there are two options.

00:46:39 Interviewee

If they do well documentation, then we can. We can read the documentation and reproduce the exact results by his own research. The other thing is if he doesn't provide then his claim should be.

00:46:53 Interviewee

According to his code, if we run that should produce automatically those results. But in many cases I still remember last year. Maybe I tried best papers from the CVPR 2020 twenty and and 19.

00:47:08 Interviewee

18 even 181718.

00:47:11 Interviewee

But I could not produce the exact same results. It was 90% failure.

00:47:18 Interviewee

And sometimes there are different kind of challenges. There are different kind of.

00:47:24 Interviewee

Hidden things. Sometimes they use legacy technologies.

00:47:29 Interviewee

You are sitting or you are using.

00:47:31 Interviewee

The latest stack.

00:47:33 Interviewee

Suppose Pytorch, I'm using Pytorch 1.6.

00:47:38 Interviewee

But that paper is producing result on 0.41.

00:47:44 Interviewee

Now that people.

00:47:46 Interviewee

That model or that either I have to transform that code and then read it in the model other I have to convert those modes under that and then I have to use. So this is the again one of the biggest challenge and we need some decisions some revolutionised.

00:48:07 Interviewee

Strategies or some good advice on that?

00:48:11 Interviewee

Area that, just like a software engineer, researchers never think like that.

00:48:17 Interviewee

To while writing a paper, because this is not their job, this is not their focus. I'm I'm totally agree on that.

00:48:25 Interviewee

So there must be some some middle position they are called search engineering positions. They must transform if that paper is really good and that is under the gas station analogy that that is useless.

00:48:38 Interviewee

If you are giving me a very good.

00:48:42 Interviewee

Ragged paper or value paper, but that is the legacy, so no one will use it. It's it will just remain our research.

00:48:50 Interviewee

But if that paper is with with little more effort and the same researcher or the same research engineer who can upgrade their code, upgrade that model, upgrade the documentation, or upgrade the things that are associated with with that.

00:49:06 Interviewee

Then definitely that research can boost a lot and can be contributed not only in the research for other people, but also for the purely industry as well.

00:49:16 Interviewee

So this is another challenge that I I feel and I found in my last three-year careers with, with discussion, because I'm continually in touch with the.

00:49:25 Interviewee

People who are researchers in my own institute, they are more than 80 people who are.

00:49:30 Interviewee

Postdoc PID candidates or researchers all over the world. So because this is going to be in the hub of the AI technologies in " " which is representing the.

00:49:44 Interviewee

Bull failure? I hope so. I think I have covered all these points. Whatever my experience, whatever things I did in this.

00:49:52 Interviewer

Yeah, yeah. Yeah, exactly.