**Interviewer 1:** [00:00:00] Okay, if you don't have any questions, I would proceed with some background questions. So could you tell us what is your current position in your job?

**Interviewee:** [00:00:11] Yeah, I'm currently working at [Company Name: Removed for Anonymity] as developer for connected vehicle cloud, and it's a cloud which supports different operations on [calls/cars?]. For example, I don't know, everything which is related to traffic everything which is related to individual calls and somewhere. So it's more of a cloud which connects all vehicles, basically, that's why it's called [inaudible].

**Interviewer 1:** [00:00:39] Okay, so can you tell us what is your overall job experience and specifically your experience in deep learning or machine learning?

**Interviewee:** [00:00:48] Okay, so my overall job experience is around 10 years this is more of development experience. It's not specifically machine learning it is more Java, golang, PHP, little Python. And yeah, some other languages for programming. For machine learning, I has started its maybe 5 years ago or something. We built first speech recognition system, I started with your speech recognition system and that time it was.. Actually I've started in 2012 and there was no deep learning during this time - deep learning arrived a little bit later. First thing which I started with was a robot motion and it was a vision task, basically. And then after that I switched to sequence modeling when deep learning arrived, and we were doing some speech recognition systems and then I was more of C++ modeling. Let's say I was researching the sequence modeling and during that time I started with context recognition problems when you have some text and they want to answer questions to the text, Q&A session, basically.

**Interviewer 1:** [00:02:15] Thank you. So could you tell us what type of deep learning networks you have developed, implemented by that I mean supervised, unsupervised, reinforcement.

**Interviewee:** [00:02:27] I haven't looked at the reinforcement learning but I've done some research in doing the first part of like vision problem. I've done PID controllers and also fuzzy neural networks, it was back [inaudible] and then, when we started with [inaudible] modeling I've used LSTM.

**Interviewer 1:** [00:02:56] Okay, thank you and which programming languages and Frameworks have you been using?

**Interviewee:** [00:03:01] Its Deeplearning4j -Java and also Python.

**Interviewer 1:** [00:03:10] Okay, have you used a frameworks like tensorflow, keras, pytorch?

**Interviewee:** [00:03:13] Not much, I used Keras little bit, but I do not like it. It's more of, you stack together everything but it's not giving you a [flexibility /possibility ?].

**Interviewer 1:** [00:03:24] I see.

**Interviewee:** [00:03:25] Of course, it's good because then you don't have to know a lot about about programming.

**Interviewer 1:** [00:03:32] You prefer to know more about programming.

**Interviewee:** [00:03:35] Yeah

**Interviewer 1:** [00:03:35] So as I tried to explain at the very beginning, we have one general question, which is what are the bugs that you faced very often and what are the problems? So if you if there is anything that comes to your mind you can tell us about that. We would like to start with that.

**Interviewee:** [00:03:54] I mean, bugs that are related to bias and variances and stuff like that. So it's more of a data problem, labeling and also like if you don't have so much data to train a network, sometimes you get some problems doing this training. And I think this is main common problem, which I experienced.

**Interviewer 1:** [00:04:18] Okay, so let's start with training data. So what you said that you you do not have enough training data very often Your solution to that is just to try to get more training data, or do you try to address it in some other way?

**Interviewee:** [00:04:36] Yeah, adding noise and stuff like that too. If you have a speech recognition system, adding noise [inaudible] the examples and stuff like that, that helps and also, I don't remember the exact term, but for a vision problem, we always change the ... [forgot the word]

**Interviewer 1:** [00:05:03] Dimensions of the picture?

**Interviewee:** [00:05:05] Not dimensions, rotation, but dimensions as well.

**Interviewer 1:** [00:05:09] Is it that for that augmentation?

**Interviewee:** [00:05:14] Yeah, exactly, data augmentation.

**Interviewer 1:** [00:05:17] You also mentioned labeling problems. Do you have to label your training data manually? Sometimes yes, but mostly you can use some unsupervised learning algorithm which can label for you, not in all examples, for me at least, but..

**Interviewee:** [00:05:35] So you mean that some training data is hard to label, right?.

**Interviewer 1:** [00:05:42] Okay, and any other problems related to labeling.

**Interviewee:** [00:05:47] Yeah, I mean there's some problems it's not programming problems. There's more use case problems, like for example, if it is speech recognition system and you have English recognition and some person uses some worlds, which are not like English words or some specific words. Basically, I mean there are words which are not English words, but there are also terminology. This kind of things is very hard to map to the context. For example, if someone uses for example, Java, that's that program language, but it's not a all real English world, I would say. Oh, in some contexts could be very confusing. For example, if you use as the Java programming language and the network doesn't understand the context of the text.

**Interviewer 1:** [00:06:46] Okay, so do you always collect your training data, or did you ever use the existing data set?

**Interviewee:** [00:06:58] Yes,I'd say most of the time I have had data beforehand. So there I haven't collected anything but some adjustments or like data augmentation or something like, that was done for me. But most of the time it was there already.

**Interviewer 1:** [00:07:24] Could you tell us if you have ever had to pre-process the training data? Do you remember any processing steps that helped you to achieve better performance or to train faster and things like that?

**Interviewee:** [00:07:38] You mean PCA or something like that or what..

**Interviewer 1:** [00:07:46] Yes, Yes, any pre-processing.

**Interviewee:** [00:07:47] Yes, I've used PCA and some pre-processing algorithms, but I don't remember exactly which ones I've used.

I remember PCA because I've used it recently but don't remember the old ones. Probably, yes, but I do not remember.

**Interviewer 1:** [00:08:08] Okay, any other kind of problems related to training data that you have faced. Does anything come to mind?

**Interviewee:** [00:08:17] Yeah computer power or something like that, but recently I've used only clouds. So it's fun, I can bring up a lot of nodes It's at one moment of time and then kill all of them.

**Interviewer 1:** [00:08:34] Okay, so using cloud improved this experience for you lets say.

**Interviewee:** [00:08:38] Yeah.

**Interviewer 1:** [00:08:39] Okay, so. Let's talk about model structures maybe. So did you ever have problems related to the wrong model structure like the type of the layers, the number of the layers?

**Interviewee:** [00:08:55] Yeah, of course, but it's more of..I don't know if It's a problem or not and maybe that's why I did not mention it. Let's say you start with something and then you see that it's not fitting I mean [inaudible] overfitting or I don't know, underfitting and then you start to add more layers depends on what the problem you are having.

**Interviewer 1:** [00:09:19] can you give us any specific example where you had to, you know add a layer?

**Interviewee:** [00:09:23] I would say any network which I did. I mean by finding what is the causing issue, I mean, by increasing decreasing data and applying some regularisation methods I was able to find what is the error and then the the decreasing or increasing layers could help it sometimes.

**Interviewer 1:** [00:09:52] Did you ever had problems with layer dimensions and things like that?

**Interviewee:** [00:09:58] Dimensions, you mean how many hidden layers I have or how many nodes in one layer I have?

**Interviewer 1:** [00:10:04] How many nodes in one layer, yes?

**Interviewee:** [00:10:06] Yeah, that's also a problem of Dropout and things like that.

**Interviewer 1:** [00:10:11] So you added Dropout layers to deal with Dimensions.

**Interviewee:** [00:10:17] Yes

**Interviewer 1:** [00:10:18] Okay, about hyperparameters. Have you ever had problems with hyperparameters tuning.

**Interviewee:** [00:10:27] Everybody are have them I think. Yes, setting them is not that easy. But yeah, I've used some like varying let's say learning rates, things like changing learning rate over time. Also I used some fixed learning rate and trained, saw the results and then I changed and trained again and stuff like that. I'm sure maybe there are better alternatives for that, but I haven't tried it.

**Interviewer 1:** [00:11:05] Okay, did you have problems related to batch size, to a number of epochs, things like that.

**Interviewee:** [00:11:14] Of course, that's another thing. I haven't trained a big Network, so I haven't really had millions millions of data. So that wasn't the case here. But yeah, I think stochastic gradient descent can help or mini-batch something. I've used that.

**Interviewer 1:** [00:11:57] Did you use the predefined or custom return loss functions in your experience? And did you face any problems related to loss function?

**Interviewee:** [00:12:05] No, I've used the standard.

**Interviewer 1:** [00:12:11] And it usually worked for you. So did you ever have to change it or..

**Interviewee:** [00:12:16] It's usually worked I haven't I mean this problem is maybe it's because I haven't had a lot of different problems which are specific.

**Interviewer 1:** [00:12:31] Okay, so you told us that now you're using Cloud but maybe before in your experience you had any hardware related problems when training models.

**Interviewee:** [00:12:43] Yes, of course when we started it was mainly CPU stuff, which was very popular at that time. Then we started to buy GPU and I mean. I mean after some time we start to buy servers, dedicated servicse from some companies. I had some powerful computers before. Then, I can just buy some servers, train on them and then canceled it.

**Nargiz:** [00:13:27] When you bought some GPUs did you try to run your code on multiple gpus at one time and did you have any problems with this like concurrency problems?

**Interviewee:** [00:13:38] [Inaudible]

**Nargiz:** [00:13:40] Could you please tell us what type of problems did your face?

**Interviewee:** [00:13:46] I had my computer. It was one GPU and there wasn't concurrency on my computer at least, but then after some time we started to buy, so I mean to buy/rent dedicated server . So not yet. I don't remember that. I haven't had any problems with concurrency.

**Interviewer 1:** [00:14:11] So you were telling us about this speech recognition implementations that you had. So for that you said that you used some existing data set, right?

**Interviewee:** [00:14:24] Yes

**Interviewer 1:** [00:14:25] So do you remember any specific steps that y taken to improve their accuracy of your model? So that it does, you know, more proper speech recognition?

**Interviewee:** [00:14:39] What do you mean [inaudible]

**Interviewer 1:** [00:14:43] like any steps that you have taken during your experience to improve the accuracy or performance of your models? Rather than hyperparameter tuning or etc.

**Interviewee:** [00:14:56] Yeah, I mean things like hyperparameter tuning, dropping some layers, checking and adding some more or less, these things, mainly.

**Interviewer 1:** [00:15:10] Okay. So other than speech recognition, could you remind me what other kind of problems you address with deep learning networks?

**Interviewee:** [00:15:20] Currently, I am working on context [inaudible]. I mean they text and you answer the questions in these texts, but..

**Interviewer 1:** [00:15:34] No, could you give more details?

**Interviewee:** [00:15:36] So, contex information. I mean there's a text and there are questions to this text and you answer it's. Q&A sessions. I'm using Squat [?] which is Stanford's question and answer system and but I haven't started with it. I mean, it's just in the process..

**Interviewer 1:** [00:15:54] Okay, this system learns from somewhere and it's supposed to answer the questions?

**Interviewee:** [00:15:59] Sorry, I didn't hear the first part.

**Interviewer 1:** [00:16:02] So your system, it learns from somewhere and then it is supposed to answer questions? Is it IBM Watson kind of thing?

**Interviewee:** [00:16:10] Yeah something like that.

**Interviewer 1:** [00:16:11] Okay, so what are the sources of learning?

**Interviewee:** [00:16:14] One is Wikipedia

**Interviewer 1:** [00:16:17] Wikipedia, okay. So then you have to process big amounts of text, right?

**Interviewee:** [00:16:22] Yes

**Interviewer 1:** [00:16:23] Maybe you could tell us what kind of problems you face when doing that.

**Interviewee:** [00:16:27] It is in the first stages right now. I haven't even started.

**Interviewer 1:** [00:16:34] A, okay, you haven't started it yet.. Okay, then you worked on image classification. No, I am making this up.

**Interviewee:** [00:16:45] Yeah, vision problem, convolutional things. Yes.

**Interviewer 1:** [00:16:47] Okay. And what kind of images were you trying to classify?

**Interviewee:** [00:16:52] More traffic related images.

**Interviewer 1:** [00:16:54] Traffic related images, okay.

**Interviewee:** [00:16:56] Like Traffic lights, other cars, stuff like that.

**Interviewer 1:** [00:17:01] Was it for like a some self-driving car system?

**Interviewee:** [00:17:05] Yeah

**Interviewer 1:** [00:17:05] Okay. That's very interesting. Where did you get your training data in that case?

**Interviewee:** [00:17:11] I will tell you one thing before I proceed with that, it's more of a commercial project. I can't give you a lot of things but be had already examples of the information because it was big company behind that. They gave us everything that we needed. Related to that, we also had problems with padding and the losing features and stuff like that. [inaudible]. If you're interestedin this, and it was very big Network and lots of layers and stuff like that. So we had some problems related to training itself because it was taking a lot of time but we had very good computersin data center. So we could handle the problem.

**Interviewer 1:** [00:18:04] So was it mostly like object recognition first to see what kind of traffic sign you have or is it the traffic light and if it's a traffic light, what color is it? Okay, and so this requires some manual labeling doesn't it?

**Interviewee:** [00:18:23] Sorry?

**Interviewer 1:** [00:18:25] So this requires a manual labeling?

**Interviewee:** [00:18:28] Yes. Mostly yes.

**Interviewer 1:** [00:18:30] Did you get the data with labels or ..

**Interviewee:** [00:18:33] Yes, we got it with it. There was actually a department that worked with labeling only.

**Interviewer 1:** [00:18:38] Okay. So did you ever had an experience of getting the wrong labels or things like that?

**Interviewee:** [00:18:45] Yes, but it was so less so it did not actually affect anything.

**Interviewer 1:** [00:18:52] Okay. So did you ever apply this recognition to the life system?What was your accuracy level when you were having it in the real world system if that's an information you can give.

**Interviewee:** [00:19:10] Yeah, actually we calculated this for the self driving cars and I can't give you the exact number.

**Interviewer 1:** [00:19:22] Okay, so, okay, maybe we can...

**Interviewee:** [00:19:25] For the speach recognition, it was around, it was five years ago. So I can't give, I do not remember that number, it was around 92 93 and it was okay for that time.

**Interviewer 1:** [00:19:41] What are the measure of the quality for speech recognition? So you have the actual text that you should get and you have what you did get, do you calculate some kind of distance.

**Interviewee:** [00:19:58] Yes, exactly. That's what I mean. We did like several tests for that. So it was, first it's like of course the technical parts F1 and stuff like that. But for the for the actual examples, it was customer service who was requesting it so people called the customer service and then it was like one computer and one person who was actually speaking and they will transfer it to text and we were actually counting what is accuracy and person received a call and what [inaudible] which gives you the text back. So here it is Real Time test I would say.

**Interviewer 1:** [00:20:50] Okay, but yeah, but what is the measure of accuracy in this case? How many words are correct?

**Interviewee:** [00:20:58] Yeah. That was the case.

**Interviewer 1:** [00:21:00] So it's number of words?

**Interviewee:** [00:21:02] Yeah,

**Interviewer 1:** [00:21:02] number of words, the context

**Interviewee:** [00:21:04] at least and also we were calculating the difference between each word. So let's say you have 10 words in one sentence. How many words machine got correctly and how many words person got correctly?

**Interviewer 1:** [00:21:22] Okay, so it's kind of a lot Levenshtein distance but in words.

**Interviewee:** [00:21:25] Yeah.

**Nargiz:** [00:21:27] You mentioned some technical bugs. Could you please tell us more about that? Because we're also interested in small technical bugs you get, like everything.

**Interviewer 1:** [00:21:38] Crashes, error messages.

**Nargiz:** [00:21:41] Yes, and the tiniest things like anything you can come up with.

**Interviewee:** [00:21:47] Maybe, out of memory, at least? I do not remember what I spoke about, bu mainly the stack overflows, out of memories, I don't know, heating of CPU, back at that time, this kind of things.

**Interviewer 1:** [00:22:09] Is usually the out of memory thing is caused by the size of training data?

**Interviewee:** [00:22:14] Yeah, mainly like that. There were several approaches to iterate, I mean, if you have it outside in file or somewhere, then these problems occur.

**Interviewer 1:** [00:22:29] So the way to deal with it is to change how you read the data?

**Interviewee:** [00:22:34] Like one thing is like that or you just need read batches of the data not like the entire file.

**Interviewer 1:** [00:22:40] Okay.

Then, any others? Like do you remember one nasty bugs that you had to work on for a long time? Some bad memories?

**Interviewee:** [00:22:59] Losing features, for example, for pictures. That was a lot of pain. When you are padding or something like that, you are losing the features.

**Interviewer 1:** [00:23:10] And how did you deal with that?

**Interviewee:** [00:23:17] I do not remember, but it was more of enriching the data which comes and then doing padding. Let's say you have the image of 5x5 Matrix and you put more numbers on the features which are dropped out during the first part.

**Interviewer 1:** [00:23:51] Okay. So for training data, do you usually use the whole data set for training? Did you have problems when you could not use the whole data set so you h to you know, reduce it somehow or it's the opposite problem when you have to add things?

**Interviewee:** [00:24:16] Yeah. Mostly I had less data.

**Interviewer 1:** [00:24:19] And you said that you use that augmentation when you add things so could you tell us about the effects of that, how you can help you to improve your accuracy? Do you see a big difference in accuracy when you do that?

**Interviewee:** [00:24:33] Not big but it at least adds some percentages to accuracy and also because the problem was ..what I was going to say.Ah, do not mind, something I forgot. Yeah, I mean there was a problem but I don't remember what I was going to say.

**Interviewer 1:** [00:25:00] [inaudible] about data augmentation, maybe that helps you. I was asking if..

**Interviewee:** [00:25:04] Ah, yes, yes, I mean sometimes when you add more pictures to data set it could be overfit so the overfitting of the network problem. So sometimes data augmentation can help, sometimes it can damage. Well, then we could just remove these pictures and see if it affects or not.

**Interviewer 1:** [00:25:27] Did you apply data augmentation when you were training for this self-driving car thing?

**Interviewee:** [00:25:33] Yeah

**Interviewer 1:** [00:25:35] What kind of operators you usually apply? Like a rotating a bit adding, playing with colors? Yeah, if you could..

**Interviewee:** [00:25:42] Yeah more of rotating, but the rotating because you cannot rotate the picture upside down. More changing the in the right or left directions.

**Interviewer 1:** [00:25:55] Okay, so. And for example the colors of..

**Interviewee:** [00:26:03] Yeah, we changed it to a grayscale and something like that. It was also the case. But I would say, more of rotating than cloring them.

**Interviewer 1:** [00:26:15] So how do you decide, let's say you have to apply data augmentation. How do you decide what exactly you should do and did you ever have cases when you made wrong decisions, but then you changed them so better ones.

**Interviewee:** [00:26:32] It was more of trial and error, lets say.

**Interviewer 1:** [00:26:38] It was more what?

**Interviewee:** [00:26:39] Trial and error. I mean, you try one thing and then you check what is output and then if it's a little less than you change it in the other direction.

**Interviewer 1:** [00:26:52] So for this system or for any other, do you remember how long usually the training takes. Is it hours, days?

**Interviewee:** [00:27:01] No, it took weeks at least for some problems which we had with self-driving cars. It could take more weeks. It depends really where we were running it because we had several environments where we could run it. So it was [inaudible] these kind of things and in each environment it could take different amount of time depends on how many machines we are having available.

**Interviewer 1:** [00:27:29] Okay, thank you. So I'm guessing you were using simulators when doing this testing and training thing. Is it? Yeah, I'm not sure if you can tell us which simulators are.

**Interviewee:** [00:27:41] No no in this project I can't even tell which tools and for whom it was a project , there is a big company behind that but I can't tell the company name. But for the vision problems, before that ,robot motions I used SimuLink before.That was one. Also, the one was Web something. WebRobot, I think.

**Interviewer 1:** [00:28:16] So again for this project were you also detecting other cars around this car or people who are crossing?

**Interviewee:** [00:28:27] Yes, pedestrians, traffic lights, other cars and other objects. Let's say you have some sort of, I don't know, in the middle of the road you have something.

**Interviewer 1:** [00:28:41] So which type of objects are the hardest to detect?

**Interviewee:** [00:28:49] The cars, maybe? because there are different shapes of cars, lets say.

**Interviewer 1:** [00:28:54] So for for detecting cars, you also had to train on different car pictures, right?

**Interviewee:** [00:29:01] Yeah. Okay, Then you actually have this, what its called, rounding this with shape. So then you find the car the object you just like, I do not remember the term, but you put the border on the object and then you just yeah.

**Interviewer 1:** [00:29:26] So in this case if you have different cars and you recognize them as objects, I'm guessing that what you recognize is not of the same size while for training you should bring them all to the same size, I would guess. Did you have any problems because of that?

**Interviewee:** [00:29:42] Then [inaudible] as well as mentioned before you take the picture and either make them all one size or just make it smaller or bigger, but mainly, because we had to have big accuracy in this case because 99% is not enough mainly and you have to change.

**Interviewer 1:** [00:30:12] Could you tell us which level of automationfor these cars you were aiming for like completely automated self-driving cars.?

**Interviewee:** [00:30:20] Yeah, the company's strategy was to make completely self-driving cars, but I mean. Still, it's in progress and we already have two cars which are ready, but we are moving towards some places.

**Interviewer 1:** [00:30:48] It's like life testing, right? They're driving and..

**Interviewee:** [00:30:53] There are testing with reduced speed.

**Interviewer 1:** [00:31:02] So maybe at this point of our conversation you have remembered some other problems that value have faced. We're like an interrogation group.

**Interviewee:** [00:31:16] I can't think of something. If I find something I will write to some document or something that and send it to you if you want.

**Nargiz:** [00:31:24] Now that would be very useful.

**Interviewer 1:** [00:31:26] Yes. If it comes to mind that we really useful. So do you have anything else to add?

**Interviewee:** [00:31:33] I got your site and it's more of crime thing. You're starting to protect or what is?

**Interviewer 1:** [00:31:43] No. No, the Precrime is just a fancy name that the project has, we don't do anything with crimes. So we are a research group and our main research is about testing deep Learning Systems.

**Interviewee:** [00:31:57] Okay

**Interviewer 1:** [00:31:57] And yeah, we have like we have a lot of directions going ,self-driving cars is something we are very interested in where you would like to test them and right now yeah, what we're trying to do is usually in testing there is this notion that you should introduce a synthetic fault to the code and see if your test suite is able to catch that, it's calle mutation testing. And these faults are supposed to be aligned with what happens in the real world. And that's what we are trying to learn - what kind of bugs people get in the real world so that maybe we can come up with the taxonomy and maybe it was some proposal of adding them - the synthetic faults to the code, something like that.

**Interviewee:** [00:32:40] Okay. Thank you. It's interesting.

**Interviewer 1:** [00:32:46] So then maybe if you remember anything you can send it to us. Thanks a lot for your time.

**Interviewee:** [00:32:53] Thank you.