**Interviewer 2:** [00:00:00] Okay, good to go.

**Interviewer 1:** [00:00:03] Okay. So my first questions would be just some background questions. First question is what is your current position at your current job?

**Interviewee:** [00:00:15] I'm a machine learning engineer.

**Interviewer 1:** [00:00:18] Okay, do you work in Industry?

**Interviewee:** [00:00:21] Yeah, industry.

**Interviewer 1:** [00:00:23] Okay, so could you tell us about your overall work experience and your experience specifically in deep learning, in terms of years?

**Interviewee:** [00:00:34] Okay. Before I graduated from the Graduate School I basically did not know a thing about deep learning and I learned deep learning after I got into the industry. So it was in my previous company and I did like 2.5 years in my previous company. Yeah, I've used deep learning mainly in natural language processing. I did some project in news summarization and images and some basic recommendation systems.

**Interviewer 1:** [00:01:25] So your experience in deep learning is two and a half years. Did I get it right?

**Interviewee:** [00:01:37] Yeah.

**Interviewer 1:** [00:01:37] And your overall work experience? Like this is just for statistics, but we have to ask.

**Interviewee:** [00:01:45] Okay. Overall experience is like five to six years.

**Interviewer 1:** [00:01:55] Okay. Thank you.

**Interviewer 2:** [00:01:57] Thank you.

**Interviewer 1:** [00:01:58] So what type of deep learning networks have you developed, implemented and by that I mean supervised and unsupervised reinforcement learning and Etc.

**Interviewee:** [00:02:08] Mostly supervised. I've tried some reinforcement learning, but it was no success. So, mostly supervised.

**Interviewer 1:** [00:02:16] Okay. Thank you. And which problems you were trying to tackle using deep learning network? Examples would be image classification speech recognition.

**Interviewee:** [00:02:29] News summarization, Image search, which is basically a representation [inaudible]. Any system and [inaudible].

**Interviewer 1:** [00:02:47] Sorry, representation of what?

**Interviewee:** [00:02:52] Representation of some some clothes image. Well, the fact is a fashion dataset, so I learned with the [inaudible] vectors about the clothes and use them to search for similar clothes.

**Interviewer 2:** [00:03:07] And you told us something about NLP?

**Interviewee:** [00:03:14] In NLP I did new summarization. I collected in Chinese, Chinese news data set of about a million Chinese news and we were clustering on it. And after clustering I generate... Is a multi document and single output sentence summarization and display those summarizations on our user interface.

**Interviewer 2:** [00:03:45] Thank you.

**Interviewer 1:** [00:03:46] Thank you and which programming languages and Frameworks have you been using so far?

**Interviewee:** [00:03:51] I mostly use python. In terms of deep learning I use Keras and Pytorch.

**Interviewer 1:** [00:04:01] Okay, thank you. So for this interview, we have one general question, which is what kind of problems bugs challenges your face while developing systems that use deep learning and for you we also would be interested on what kind of questions you think people ask very often on Stackoverflow, what type of bugs to they report there so we could start from here and you could tell us whatever comes to your mind about this. That would be great.

**Interviewer 2:** [00:04:29] Yes and by bugs we mean everything - even the small technical issues you face, error messages, like everything from the big conceptual challenges to the tiniest technical bugs your faced.

**Interviewee:** [00:04:48] Most of the issues are I've encountered was in Keras. Keras LSTM or in general recurrent neural networks. The recurrent neural networks.. In the beginning like in some previous version there was a huge refactor. So after the refactoring, some bugs just came out and at the time I was working LSTM in Keras.I fixed some of these bugs and submitted to Keras as [inaudible].

Most of [inaudible] are just some quite basic like indexing error, like some layers are not indexed properly. So if you combine those LSTM [inaudible] if you stack them in a certain order the [inaudible] broke. Yeah. Some other bugs were in Pytorch, I use the openMNT, the machine translation library to implement my summarization system and the openMNT system at the time I was using it it was kind of half-finished if - it was incomplete. So it has.. The most buggy the place I've encounter is the beam search. And when I was using the beam search it was like a huge bunch of code and not very well factorized. So there were so many bugs in it and specifically, the bug is not so easy to discover. ike the output of the model still look like some valid Chinese sentence, but the score is not correct.

At the time there were two versions of beam search: a slower version and a vectorized version - the GPU version. The vectorized version don't output the same scores as non-vectorized version, but that non-vectorised version is often much better, much better [inaudible] the results are more plausible.

So I was trying to somehow fix the fast version - the vectorised version and make the both versions out to the same.

**Interviewer 1:** [00:07:41] Okay. So these are the bugs that were in the framework that you have been using, right?

**Interviewee:** [00:07:48] Yeah

**Interviewer 1:** [00:07:49] What about the kind of problems that appear because of the errors that you make as a developer? Could you tell us about those ones?

**Interviewee:** [00:08:06] The errors.. Oh, yeah, there was one time but it's not a deep learning library, so.. But in general I was using LightGBM. And [inaudible] it and after the software upgrade if I upgraded the package it just overfitted seriously.

**Interviewer 1:** [00:08:32] So what was the package? Sorry, I did not hear that.

**Interviewee:** [00:08:37] LightGBM.

**Interviewer 1:** [00:08:41] Okay.

**Interviewee:** [00:08:43] There was a [inaudible] change in the software and I didn't [inaudible] out and I just upgraded and my model just broke.

**Interviewer 1:** [00:08:51] Okay, I see. Yeah, thank you. So about Stackoverflow. So we have seen that you have a lot of points as you reply to the questions a lot. What kind of problems people usually ask about on StackOverflow? Is there like any categories that you think that appear too much, too often?

**Interviewee:** [00:09:22] There's one type of problem. I think it occurs very often but it's not easy to answer is like why do you mind model diverge or why the lost become another number and it's really hard to answer because there's often not enough details to figure out what is exactly the problem.

**Interviewer 1:** [00:09:49] And the questions that are easy to answer for you. Do you think they're all like these beginner kind of questions that appear also very often.

**Interviewee:** [00:10:02] The beginner questions or...

**Interviewer 1:** [00:10:05] No, no any questions, even if they are easy for you to answer and you don't see them as a big problem, but people keep asking about it on Stackoverflow. Do you think there's a category of questions there?

**Interviewer 2:** [00:10:16] Things like they do something not correctly because they don't have enough experience.

One thing is the dimension is match problem for example, the LSTM or convolution network? If you stack [inaudible] convolutional network and you don't understand what the parameter.. What the meaning of the parameters is, it is easy to get in some dimensions that just don't match and the model won't compile. Yeah.

**Interviewer 1:** [00:10:48] Okay, anything else?

**Interviewee:** [00:10:52] And the recurrent Network.. Many people ask about how would you implement a recurrent neural network that does not have a fixed number of time steps.

**Interviewer 1:** [00:11:12] Okay.

**Interviewee:** [00:11:13] Yeah, I think I saw these questions quite sometimes but I think it's actually in the tutorial or examples in the Keras folder. So the answer is mostly [inaudible] point to the example code.

**Interviewer 1:** [00:11:33] Okay, so, as a person who develop machine learning systems and also you answer on Stackoverflow, do you think that the problems that occur in real world and the problems people ask on Stackoverflow, are they like representative or do you think it is different/easier?

**Interviewee:** [00:11:59] I think it's somehow different like questions my what my colleagues will ask me are different from what those on Stackoverflow. I think it's because most machine learning engineers in the industry have more experience in implementing, in tracing the code. So the questions are like more deeply into the library. It's not just like, how how should I stack these layers to make a valid model but most questions on StackOverflow are like model building or why does it diverge the kind of questions.

**Interviewer 1:** [00:12:56] Okay. Thank you.

**Interviewee:** [00:12:57] So mostly it is not bug questions on StackOverflow.

**Interviewer 1:** [00:13:01] Okay, thank you. So you told us that you have collected this huge database of news in Chinese, so could you tell us anything about problems that are related to collecting this kind of training data?

**Interviewee:** [00:13:22] In terms of collecting the data I think the biggest problem is the website - they keep changing the format of UI so if you want the [inaudible]of the news title or news content you have to regularly check the website and make sure they don't change. Okay? Yeah, and another problem is that.. But this is kind of a system problem is that if you [inaudible] wrong data processing program. You typically want to run it in the multithread, I mean in parallel , but sometimes there will be some like race condition or the model just don't [inaudible] just don't behave like you think it should be like if you have 32 cores, it just don't go 32 times faster. Yeah, so you need to dive into the code and figure out what caused the bottleneck.

**Interviewer 1:** [00:14:36] And what causes the bottleneck usually?

**Interviewee:** [00:14:40] For example, some libraries like...Some libraries there are like compressed, yes, they're compressed. So the most of the computation is like decompressing the file. If load it [inaudible] decompressing or like if you try to load the data in Json format and you're spend [inaudible] a lot of computational like decoding the Json instead of processing the data. So yeah, so sometimes you have to change the format to make it really parallel. Yeah.

Okay. So after collecting these data in order for it to be used for training, did you ever have to process these collected data in any way and do you remember like the cases when before pre-processing things were not working and afterwards it kind of worked?

It's not this problem, but in another program I've tried this self-normalizing network and it requires a special normalizing initialized. Here, you have to standardize the dataset. Otherwise, it won't work. Yeah, and yeah, so in the beginning I was just trying the self-normalizing network and it just don't go as good as they really the [inaudible] network. So I was figuring out why is it like that and one of my colleagues just come here and say you need to normalize them like yeah.

**Interviewer 1:** [00:16:33] Did you ever use some existing datasets or did you always collect your own data?

**Interviewee:** [00:16:41] Well, I've used the fashion dataset, the dataset about clothes is provided by, I think, a university in Hong Kong or..

**Interviewer 1:** [00:16:53] Yeah, I think I know about that one and did you ever had problems when you were using the existing data set?

**Interviewee:** [00:17:03] Yeah, the image data set is not very accurately labeled, so it contains some kind of noise and the image domain is... This is kind of mix domain problem is not [inaudible] like because it was clothes. But there's some of the clothes images are product images. Like there are no people in the image, in some they have people wearing the clothes and take a picture. If you mix both images together the model will likely not learn very well. So we have to do some processing to filter out the images that don't have people in it and so in my case, the search result were much better.

**Interviewer 1:** [00:18:05] Okay. So do you remember any problems related to training data that you have encountered, but that I mean, like if you did not have enough training data, if it was in some way of wrong, and etc.

**Interviewee:** [00:18:24] The news summarization. Summarization requires quite a lot of data. [inaudible] it doesn't work very well, I just keep growing the dataset size and till i get like a million that scale and it's just starts working.

**Interviewer 1:** [00:18:46] So do you remember any difference in accuracy? Like how accuracy was improving as you were improving the size of your dataset?

**Interviewee:** [00:19:01] I forgot about the numbers.

**Interviewer 1:** [00:19:03] Yeah, that's fine. Also, out of curiosity, so news summarization, if I understand correctly, you get news from some websites. Do you have to write their summaries yourself so that you train the network or how does that work? Like, what does the network do?

**Interviewee:** [00:19:23] I just use the titles. I make the neural network try to learn to interpret titles.

**Interviewer 1:** [00:19:29] So it's basically predicting title for a given news, right?

**Interviewee:** [00:19:35] Yeah.

**Interviewer 1:** [00:19:36] Okay, that sounds interesting. So in this case, you don't have to do any manual labeling or anything, right? You just use the title.

**Interviewee:** [00:19:50] Yeah.

**Interviewer 1:** [00:19:50] How do you determine whether the predicted one is close to the real one in terms that its like text comparison, right?

**Interviewee:** [00:20:05] Excuse me. Can you..?

**Interviewer 1:** [00:20:06] So let's say you have this news and you have its original title and then your model predicts some other title. So it might be either close or very far away from what you have, right? So, what is your metric to determine whether this text is actually the right title or not? Because it should not be the exact match, no?

**Interviewee:** [00:20:29] No. Yeah,I use human labeling. Like I pick a hundred news clusters and just yeah just average the correct numbers. Yeah, I use the accuracy as my metric. So it is human labeling.

**Interviewer 1:** [00:20:54] Did you have like humans involved who were labeling the news for you?

**Interviewee:** [00:21:00] No, I just asked my team members.

**Interviewer 1:** [00:21:03] Okay, so I yeah, I'm asking because I wanted to ask about the agreement rate between them do they often give the same labels or is there some mislabeling going on?

**Interviewee:** [00:21:15] No, I think... because it's just a correct and incorrect decision, so..

**Interviewer 1:** [00:21:21] Okay, I see, okay. So this is what you do on the outputs. If you get a prediction, you ask a human to decide whether this one is correct. Oh, okay, I get it.

**Interviewee:** [00:21:30] Yeah.

**Interviewer 1:** [00:21:30] Thank you. So I wanted to ask about model structure and by that I mean layersm number of layers, and etcetera. Did you ever had problems related to wrong model structure where something was wrong you changed something about the structure and then it work?

**Interviewee:** [00:21:47] I think it is in the image search problem. And in the beginning I just used the open dataset and it contains labels about the category of the clothing. Like is it a dress is it a T-shirt or so. [inaudible] not enough just using those labels to create an image search system. So later I changed the architecture to include some additional loss terms, some auxiliary loss.

**Interviewer 1:** [00:22:27] Okay.

**Interviewee:** [00:22:29] Yeah, in my case the color.. When I introduced some lost term about the color it just worked very well.

**Interviewer 1:** [00:22:40] I s it like a feature that you have added or is it different?

**Interviewee:** [00:22:46] No, no, no. It is like, you know, additional output. Originally [inaudible] the code clothing category and I added another about the color of the clothes.

**Interviewer 1:** [00:23:01] Okay, so you have like multiple labels for one input, right?

**Interviewee:** [00:23:11] Yeah, yeah.

**Interviewer 1:** [00:23:11] And now your model also predicts the color or it just classifies..

**Interviewee:** [00:23:17] Yes, it is the colors of the learned features will [inaudible] about the colors. So this search results were much better.

Okay, thank you. So, about the model structure in terms of layer and their interaction with each other. Did you ever encounter any problems because of that.

Excuse me. Can you repeat it again?

**Interviewer 1:** [00:23:42] Yeah, about model structure in terms of layers [inaudible] any problems related to that?

**Interviewee:** [00:23:55] Sorry, I think it is the network connection..

**Interviewer 2:** [00:23:58] Yes, something.. So she asks, when you deal with the network structure when you connect the layers, stack them on each other. Did you have any problems related to the the connection of layers between each other?

**Interviewee:** [00:24:16] Connection of layers...

**Interviewer 1:** [00:24:19] [inaudible]

**Interviewer 2:** [00:24:21] We don't hear you, Interviewer 1..

**Interviewee:** [00:24:33] The connection of layers...

**Interviewer 2:** [00:24:38] Maybe something with dimensionality of layers - the number of nodes in them? Oh, we've lost her, sorry.

**Interviewee:** [00:24:49] So I just continue to answer question?

**Interviewer 2:** [00:24:58] Yeah, one second. I'm going to try to add her again. Sorry. I guess she has some connection problems.

Hello, are you here?

**Interviewer 1:** [00:25:45] Do you hear me now?

**Interviewer 2:** [00:25:47] Not really good so.. I can hear you but your voice is trembling.

Are you here?

**Interviewer 1:** [00:26:20] Hello..[inaudible]

**Interviewer 2:** [00:26:23] It's still not really good.

**Interviewer 1:** [00:26:30] You don't Hear me at all?.

**Interviewer 2:** [00:26:34] We can hear but not really well, maybe you can switch off the video.

**Interviewer 1:** [00:26:39] Okay, now?

**Interviewer 2:** [00:26:41] It's better I think.

**Interviewer 1:** [00:26:44] Ok, great, so let's continue.

**Interviewee:** [00:26:47] Okay.

**Interviewer 1:** [00:26:48] Where did we stop?

**Interviewer 2:** [00:26:51] You had a question about stacking layers on top of each other.

**Interviewer 1:** [00:27:05] Yeah, yeah.

**Interviewee:** [00:27:05] It's kind of.. When I was experimenting about convolutional neural network on news summarization, and I found it quite somewhat difficult to select the right number of layers. I mean it's not like.. When I use the same number of parameters in the paper, and I don't get as good results as in the paper. So, yeah.. So eventually I just decided to experiment a smaller model and a huge model and compare the number of parameters to my LTSM model. As long as they are same so I think it's a fair comparison. Yeah.

**Interviewer 2:** [00:27:54] And was it only the question of number of layers or did you also experiment with the type of layers?

**Interviewee:** [00:28:07] Mostly I experiment about the number of layers and the number of neurons.

**Interviewer 2:** [00:28:12] Number of neurons as well. Okay.

**Interviewee:** [00:28:14] I am keeping the number of parameters the same. Yeah.

**Interviewer 2:** [00:28:19] Thank you.

**Interviewer 1:** [00:28:19] [inaudible]

**Interviewer 2:** [00:28:20] Interviewer 1?

**Interviewer 1:** [00:28:20] Do you hear me?

**Interviewer 2:** [00:28:25] We hear you but your voice gets distorted.

**Interviewer 1:** [00:28:33] Okay, okay, so you go on.

**Interviewer 2:** [00:28:39] You can try to because like I can understand what you say, but..

**Interviewer 1:** [00:28:43] Okay, so I wanted to ask if you ever had problems related to hyperparameters tuning? And when you changed values of hyperparameters like increased or decreased and that improved the performance of your model?

**Interviewer 2:** [00:29:03] Did you understand what she asked?

**Interviewee:** [00:29:06] No, I can't hear her..

**Interviewer 2:** [00:29:07] So, okay. So she asked if you have ever faced any problems related to hyperparameters tuning, by that we mean batch size, activation function, learning rate and all sorts of hyperparameters. And did you face any cases when you changed the value of hyper parameter like decreased or increased it and you got some increase in performance.

**Interviewee:** [00:29:37] Yeah, I've tried to tune the..I always tune the learning rate. I think the learning rate is quite important and it affect the performance, not always very much but.. If you change the learning rate, you'll always get some improvement, at least,I think..

**Interviewer 2:** [00:30:00] Thank you and maybe anything related to activation function? Maybe you changed the type of activation function and you got some improvement in the model?

**Interviewee:** [00:30:16] I've tried changing the 'relu' you activation to the hyperbolic tangent. In my experience It didn't affect the performance very much.

**Interviewer 2:** [00:30:31] Okay, thank you. So the next question will be: do you do usually use a predefined - existing loss functions or have you ever tried to write a custom one? And did you have any problems because of the loss functions?

**Interviewee:** [00:30:55] Yeah, I try to implement custom loss function quite often like in the image search I have several different loss functions. The image category is a standard cross-entropy loss. But the color loss is a [inaudible] these things. Yeah, so that requires some custom loss function.

**Interviewer 2:** [00:31:40] Did you ever have any problems with this any bugs that you made during the writing this custom loss function?

**Interviewee:** [00:31:53] Yeah, I think the most...One of the problem I've encountered is when I will try to serialize the model and load it back. I think some previous versions of Keras is not so convenient. There are some variables that just can't be serialised very well. So you build a model and load the model weights instead of serializing the whole model object.

**Interviewer 2:** [00:32:29] Okay, Thank you. And where do you usually train your models? Do you use some specific hardware and did you ever had any problems with the hardware?

**Interviewee:** [00:32:53] In our company, we usually train models on the Amazon - AWS. Yeah, and sometimes the GPU on the AWS instance just just broke and you have to email them to fix it.

**Interviewer 2:** [00:33:14] So you don't have any of servers of your own.

**Interviewee:** [00:33:21] Yet I have a GPI you but mostly I train on the cloud service.

**Interviewer 2:** [00:33:28] Okay, thank you. And one little question when we were discussing the differences between problems you face at your workplace and ones you see on stackOverflow, you told that the questions your colleagues usually ask you are quite different from those you get on stackOverflow. Could you please give some more details about questions your colleagues asked you and which problems you usually face?

**Interviewee:** [00:34:06] Mostly my colleagues.. Most of the problms they ask are about how to implement the model in the paper. That's very standard. I mean not like the models will appear in the Keras example folder. Yeah, there may be some additional loss functions or some special structures, some special layers that are not built-in in Keras. So we try to just figure out together. Yeah.

**Interviewer 2:** [00:34:45] Okay, I see. Thank you.

**Interviewer 1:** [00:34:48] Do you hear me?

**Interviewer 2:** [00:34:51] Yes.

**Interviewee:** [00:34:51] Yeah.

**Interviewer 1:** [00:34:52] I wanted to ask the classic question of were there any very nasty bugs that you had, like the most interesting bug that you had that took a lot of your time to fix and which is related to some deep Learning System.

**Interviewee:** [00:35:08] I'm sorry, the sound i hear is...

**Interviewer 2:** [00:35:10] Ok, I will translate, so she's asking:d can you remember the most nasty bug you face in your experience of developing systems based on deep learning? Like the one that took you the most time to solve and...

**Interviewee:** [00:35:28] Yeah, I think it's still the openMNT beam search. Yeah, the new summarization search system. [inaudible] Yes. It took me quite a long time to debug it. And yeah, I think it's because the beam search itself is quite a complicated function, so.. [inaudible] Also because it's a vectorized version. So everything you printout is like huge tensor of integers. So you just don't see like nature language. It's examples so it's kind of hard to just Identified about by looking at the tensors.

**Interviewer 2:** [00:36:24] I see, and usually when you, of course you identify the bug, and usually what it is?

**Interviewee:** [00:36:35] Excuse me?

**Interviewer 2:** [00:36:36] And which bugs are the most usual ones with this system?

**Interviewee:** [00:36:50] The usual ones..

**Interviewer 2:** [00:36:57] It's okay if you don't remember. And you mentioned some tensors, so which problems with tensors you had in your experience?

**Interviewee:** [00:37:18] Oh, the reshaping, the tensor reshaping. You're trying to flatten a tensor. I mean if you try to flatten a tensor the order is very important.

Sorry, what's important? I didn't hear you..

The order, for example if you flatten the 32 times ten metrics into a 320 element [inaudible] is different from if you flatten it to 10 times 32 Matrix. The order is important and is very easy to just get it wrong.

**Interviewer 2:** [00:38:11] Okay, as I understood It happens like you this way, you occasionally mixe the order and you got some problem and you then fix it.

**Interviewee:** [00:38:27] Yeah, I think.

**Interviewer 2:** [00:38:30] Thank you and a bit of a general question. So for example, you developed some model you test it now and you see that its performance is quite good, okay, but you feel that you can improve it and you or your colleagues what steps do you usually perform in order to improve your model and increase the accuracy or whatever the output you want to get?

**Interviewee:** [00:38:59] Usually I take a look at the wrong samples. Some error analysis..

**Interviewer 2:** [00:39:08] Sorry, what?

**Interviewee:** [00:39:09] Some error analysis. Yeah, I just print out the examples and the wrong examples. So and see like.. For example in summarization why does the model generate the incorrect word. I also print out the attention - the attention scores and figure out if the model is indeed looking at the wrong position or it's some bug.

**Interviewer 2:** [00:39:44] Thank you. Can you remember anything else? Like the bug that you do too often? Maybe not too often, but if you.. Like when you are working on your models and maybe just by occasion you do something wrong and introduce some bug in your program. What is it? Even if it's something mechanical.

**Interviewee:** [00:40:15] I think there was one time I when I was implementing the convolutional sequence to sequence network and I got the positional encoding wrong. Yeah, the positional encoding. Yeah, I think it was just formulized wrong. The formula is incorrect. So. The encoding become basically effectless, so..

**Interviewer 2:** [00:40:47] Okay, thank you. And if we return to my previous question about improving your model, you told that you look at the errors and you see what was the problem and usually.. What you usually do? So after checking the problems with your wrong output, what's usually you do to fix it?

**Interviewee:** [00:41:11] Depends on the question, but one time when I experienced.. When I was working [inaudible]

**Interviewer 2:** [00:41:35] Sorry, could you please repeat because there was some problem with noise?

**Interviewee:** [00:41:39] It is one time when I was working on the new summarization. I tried to print out the attention weights and I figured out it is some kind of error in tokenization. So the model just copies the long token from the text to output but the tokenization itself is wrong. So this is the problem in Chinese processing we have to do tokenize all of our words. Yeah, so I fixed it by changing to another segmentation system. Originally I was using some kind of just Chinese segmentation and I switched to the library called sentence piece. It was some Library based on like.. It used some kind of mixed character and word tokenization, [inaudible].

**Interviewer 2:** [00:42:58] Okay. Thank you. So if you have anything to add or if you remembered any other bugs that we didn't discuss today.

**Interviewee:** [00:43:23] Yeah, there's one bug.. there is another bug about parallelization doing representation learning on a huge graph. When I was working on not so large graph like a hundred thousand of notes and a couple of edges then it works fine. But when I increase the size of the graph to like 10 million and the algorithm just freezes. Yeah, I can.. I spawned a couple of processes but the parallel processes it just don't.. Don't terminate. The entire program just freezes and yeah..

**Interviewer 2:** [00:44:19] Did you get any solution to this?

**Interviewee:** [00:44:22] Yeah, I think it's kind of problem in the python multi-processing and I solved it with.. The problem occurs whenever you need to send a huge result back to the main process - from the child process to the main process. If the result is too big the program just freeze. But if you send the result back in batch, the program will just end.

**Interviewer 2:** [00:44:57] Okay, I see, thank you and the very very last question. Sorry, we are a bit long today. So maybe you can remember something interesting you've seen on stackOverflow. Something not so really typical and not like the beginners question when you have to point the person to the tutorial, but some interesting bug you've seen there which took some of your time.

**Interviewee:** [00:45:30] Is it specific to bugs or all type of implementation questions?

**Interviewer 2:** [00:45:36] No, we are more interested in bugs and problems because the questions like 'how to do this' are not really a bug, the person just doesn't know.. So we're more interesting in problems people face and in some bugs they do.

**Interviewee:** [00:45:52] Okay. There was a bug, I think it's like why does my network output not a number? Yeah, whenever they're using the 'elu' activation. Yeah, and I tried to figure out and it is a problem in the tensorflow implementation of the activation function so it is in the C++ code is not in the Python API. So yeah, the way tensorflow implements it somehow with will generate a..I think it's a divide by zero or something, but it's a number that will cause the output to become not a number..

**Interviewer 2:** [00:46:54] So basically Tensorflow had some bug in their implementation. Or did they just do it in a way that user didn't expect?

**Interviewee:** [00:47:07] Yeah, they do it in a way that's not so safe numerically.

**Interviewer 2:** [00:47:12] Okay. Okay. Thank you. So, I guess that's it for today. Thank you very much for your time.

**Interviewer 1:** [00:47:21] Thank you, that's great. Thanks a lot, sorry I disappeared.

**Interviewee:** [00:47:27] Okay.

**Interviewer 2:** [00:47:28] Thank you very much, have a nice day.

**Interviewer 1:** [00:47:29] Bye.

**Interviewee:** [00:47:29] Okay. Bye. Bye.