Interview 9 - PO

Interviewee

yeah, I perfect. I did.

Interviewer 1

Thanks. Um, so I'll go over a quick description of the interview. Uh, so what we want to do is a, quite the

catalog of quality issues and machine learning software system. And what is the learning software system?

It's any, it's just a software system with a learning component and you can think of a recommended system,

for example.

Interviewer 1

Um, and yeah, that's pretty much it. We will ask to you around 20 question. And, um, yeah. Do you have any

Interviewee

question? Uh, yes. Well, um, do you want introduce myself or, yeah, that was the

Interviewer 1

next question.

Interviewee

Yeah, go for it. Okay. So that was the first question. Yeah.

Interviewer 1

So, um, uh, what is, can you give some background information about you?

Interviewer 1

Uh, what is your current position, how much experience you have in the

Interviewee

master learning? Sure. So Iist mainly just my, my back and I have a master degree in theoretical physics in, in

the, in Country 1, from University 1 and. I have almost seven years working in data science environment. Uh,

mainly as a data scientist. Um, I started, um, working, um, as analyst and consultant in the wind energy, um,

domain by making predictions and forecasting failures in large, uh, wind farms.

Interviewee

Um, After that I work it in a, in the financial tech sector with, um, some company that doesn't exist anymore,

but his name was, uh, Company 1. So the objective of this was to create, uh, investment portfolios, uh, based

on their risk behavior of the people. And after that, I work it almost two years in a company called Company 2.

Interviewee

There is kind of a cool in Continent 1, they do a tech education and technology. So a large, uh, amount of

courses for. Society. Uh, here I participate in projects as search engine recommend their as systems. Um,

show

Interviewee

modeling. After that I work it in, uh, global, uh, large software company, um, from Continent 1 nowadays

based, and in our states, it has more than 5,000 employees in, in.

Interviewee

The world, uh, here I work it as a data, senior data scientist, software designer, and mainly, uh, developing

NLP models. So any all, um, that, um, models related with natural language, uh, processing mainly. Um, and

after that I joined a team of in which I have. Work it for the last, uh, year, especially, and here I have

participated in several projects, forecasting, NLP cetera.

Interviewee

That's great.

Interviewer 1

Thank you. Um, what are the main qualit issues you have encountered with your data model or systems

Interviewee

so far? So. Quality issues in data and model. And basically those two fields.

Interviewer 1

Sorry, what was

Interviewee

the question? So can you, can you repeat the question again?

Interviewer 1

Yeah. What are the main quality issues? Uh, you have encounter with data more the model or the system?

Interviewer 1

The me learning so forth. The

Interviewee

system well related with data often as a data scientist, um, databases are, are doesn't have, um, always is easy to find like a holes or new values. Um, that's something that happens often. My, my point of view is most of the data basis of the work were not created with the objective of made analytical a.

Interviewee

So databases, uh, transactional databases just keep data from users or, and most of the cases, the data is over righted by new data because was not the main objective to keep, um, historical track of, um, of the data. So, um, and I see this problem in, in several industries in which I have participated in. So from the data quality.

Interviewee

It's very rare to find a company that thinks in a very analytical way or, or have a very, uh, good, uh, data culture develop it. Normally the data science, when they start in the startups or, or small companies, they start to use like the databases that the backend is using or, or the front end that are not rigorous designing to do analytical a.

MLP-artefact-management-issues

Interviewee

So that's one kind of the problems that I often, uh, found and well, cause I have participated in some data culture. Uh, can I say methodologies in, in companies? Um, with the time the companies get like, um, It, it can, what database of the companies start to be improved? So concepts as that are hou or things like that, I think a lot of companies are missing, but there is good to know that most of them start to be worried about it and, and improve this about the models.

Interviewee

Um, I think that to have, or, or design. Pipeline end to end is, is very difficult to the company and takes a lot of time to, to arrive to that point. And between start from zero and get a complete model deployed working. Um, it takes too much time and a lot of experimentation during the process. So it is very hard to arrive to that moment and, and sometimes it takes a lot of experience or months.

Interviewee

And I consider that that right now, there is a lot of tools that can, um, increase the. The performance when you are trying to, to create a pipeline. So for example, Kedro uh, is one of the frameworks that I like much more to, to work. Uh, we are using tool X for one project and an as this amazing whole, the entire workflow pipeline of the data pre processing, uh, manipulation and model and version is, uh, being a story to created and manage.

MLP-pipeline-jungle-fix

Interviewee

So again, I think those are concepts, concepts that are still evolving. Mm. But more, uh, also there is like, uh, some, well, most of the data scientists, when they start projects, they want to do, uh, complex models and normally business doesn't require code with complex machine learning models. So that. The time can impact the quality of the solution when you have too much parameters, because it's very hard to analyze why the model is behaving in some way.

Interviewee

But, uh, yes it is. I think that there is, um, on how can I say that?

Interviewee

Let me see under valuation of, uh, some machine learning models as the decision trees or even the logistic

ion. I think there is a lot of power in just simple solutions that can bring a lot of value to, to, to business. So

again, this is a matter of KU. I will say people like to do complex things when it's not.

Interviewer 1

Yes, that's for sure. Yeah. Uh, that's very interesting. Thank you. Uh, I would like to dig a little bit deeper in

one of the, uh, points you mentioned. Uh, so you talked about Kedro uh, and you said it makes deploying ML

pipeline, uh, easier. Um, what is the problem that you have when you don't have Kedro why is it difficult to

deploy a

Interviewee

ML pipeline?

Interviewee

What, well, um, Normally, well, a project data science requires like a several steps that are not well defining at

the beginning. When you start a project, I think, um, well I always say that data science is not a. As work as

author projects. So for example, a scrum can work for backend development or front end development

because you have a very clear objectives, but when you are doing data science, you are doing science and

science is basic on experiment that are basic again in hypothesis, you are, so you are testing hypothesis and

you don't know if your objective will be achieved or not.

Interviewee

So. All the process that you are constructing, maybe it will not work. And, and, and that requires a lot of, um,

research at the beginning, uh, data pre processing data manipulation, and often that takes more than the 50

or 60% of the project create a model in a project is a very short part of them. Once you have, uh, um, pre

processed your database and created a lot of features.

ML-pipeline-issues [MLP]   MLP-pipeline-jungle

Interviewee

Arrive to that point to have a very, um, clean database to use in a model is, is very hard. And.

ML-pipeline-issues [MLP]   MLP-pipeline-jungle

Interviewee

Normally, if you are not a very well order person, you will start to get a lot of repositories functions and can be a disaster at the end and notebooks. So Kro in some way is, is very robust and facilitate the way that you will conserve this pipeline step by your step until you reach, uh, very, uh, clean data.

ML-pipeline-issues [MLP]   MLP-pipeline-jungle

Interviewee

So, uh, once we discovered this in, in the project, I, I say to myself, Hey, I would like to have this in previous projects because it was very organized the way that we created the, the pipeline of pre processing and even the, the modeling, um, because it's integrated with ML flow. It was very easy to, to keep track and version the models even.

Interviewee

And. But now it's very, it is just ready to be deployed. So I, I think that the tool is what's hard at the beginning, but now it saves a lot of time when, when you are thinking all the process that we, we have, uh, all the. That we have done in the previous month. It is like, wow. Amazingly looks very nice and beautiful.

Interviewee

Um, I think about that because in previous projects, in which I work in the past, never work as fast as this one. That's why I like it. This, um, framework in general.

Interviewer 1

Okay, great. Thank you. All right. Um, so I, I will ask you questions about data collect. um, so did you ever use, did you ever leverage the services of data collectors, so people that manually collect data for you, uh, to train a model?

Interviewee

Um, no. normally the declaration is well, There is services, databases. I didn't participate too much and create something to together data from, from people. Well, from, from some, uh, tool that we have developing in, in one project, the search will start to, to record the search, um, the search patterns or words that the people

were using, uh, in the webpage used to start to discover what the people was interested for.

Interviewee

But. Here the deposit was us purely to the backend team to, Hey, I need to start to store some variables. Please create a database for me. Um, or maybe forms sometimes you ask to create forms for the people, but very simple. Okay.

Interviewer 1

Yeah, no, it's alright. Um, it. It's normal that some survey you, you don't necessarily have some experience.

Interviewer 1

We are just probing for different, uh, prob probable experience. So, yeah, no worries. Um, did you ever use external data? Well collected external. And what I mean by external data is public data set a third party, API web script data,

Interviewee

anything I remember for one financial project, I. I wanted to find a correlation between the social economical status of the people related with their incomes.

Interviewee

So I just made some web scraping, uh, of webpages of real estate to have an idea about how the, the values of the, of the distribution of the, of the price of houses was made in, in some sections and correlate with. With some other information that we have of the class. So location by the IP address C and in other project in which I work it in, in , uh, we wanted to design, um, some models that auto complete, uh, code when, or enhance the production of, uh, developers.

Data-collection-external   Data-collection-web-scrape

Interviewee

So we started to make, uh, language models and the database was the entire GitHub. I was. The person that, um, made the distraction of the database, but, uh, someone made the whole distraction of the, all the open source, uh, repositories. And we started to use that for, for the project. Okay,

Data-collection-external   Data-collection-web-scrape

Interviewer 1

great. And did you ever add the quality shoes with these data?

Data-collection-external   Data-collection-web-scrape

Interviewee

Uh, sources? Well, yes, because. Um, okay. And, and just to add other things, normally you need external data. Um, so for example, um, well metrics of, uh, economy metrics that are always changing during what the evolution of the work. So those are very easy to, to track tracking some open source, uh, database, um, quality issues.

Data-collection-external   Data-collection-web-scrape

Interviewee

Um, For the, that one that is related with real estate. I remember there is a lot of cases where the data doesn't have any sense, but maybe this is because the people is their responsible to create their, their, their data sets. And it's easy to them to, to make mistakes like, uh, instead of art for bathrooms to.

Data-collection-external   Data-collection-web-scrape

Interviewee

Particular how they can put fourt. So you, you start to see outliers in, in the dataset, um, for that one that comes from heat help. Um, I think someone make a very good, uh, cleaning process and try to select like the best repositories, because the idea was to, to work with the functions with a very clean version and.

Interviewee

So, yes, you can find quality issues all the time, because there is no regular patterns you will find outliers or things that doesn't behave as, as you expected because there are errors.

Interviewer 1

Yes. Okay. And did you ever use data generated by another system? So it can be ML based system or not. It could be any kind of system.

Interviewee

Uh, once I remember a project we would wanted to MultiTech, um, and they think it's a very unbalanced project, unbalanced data set. So we use with a library that can generate some, some data between the, the class, the minority class. So I used to have methodology, but, uh, their lessons was. Too much. Uh, well, the at then the difference was not too much to be considerate, so yes, I use it, but I, in my experience, I didn't see like, uh, increasing of the performance or so I, I mean, I always want to work with, on their ballot, on their yes.

Data-collection-system   Data-imbalanced-dataset

Interviewee

On their balance. Methodologies. I think it works very well. Well, it depends on the case how much data you have to do that process. So for example, is, is a case where we are working with under and because we have a lot of data we don't require to or case the data. Um, let me think if for language models I have.

Data-collection-system

Interviewee

A particular case. Okay.

Interviewer 1

Thank you. Um, have you ever measured the quality of your data and, or tried to improve it? I, I, I suppose, yes. From what you said? Uh,

Interviewee

yes. Yes. I think you, you know, I, from, from the last two years, my experience is that work more on the data that in the. Work on the model, you can increase the performance, but your gain can be 1%, 2%.

Interviewee

But when you work on the data and you create new features or you start to dig more and see what, why the are being closer, the increase on the performance of the model goes very high. Like you can increase your performance by 5%, 6% or, or more than just playing with hyper parameters. So I believe that they're more over by trying to clean your data prepared through the data, or as you say, like, uh, uh, generating new data from the same data.

Interviewee

Um, so, well, not didn't bring very good vessels when I test it. I don't know who, what I can do. In fact, this now. But often when you see computer computer vision, you see the, an image. You can do a lot of things to, to generate new data and improve your model. Like the

Interviewee

solution,

Interviewee

language models have some kind of, of things that you can do. Like it changed some words by C names that helps that can help the model. But, um, so, so yes, I think, but from my point of view it accelerates their process. So yes, that's where it is. Okay,

Interviewer 1

great. And do, do you have any tools and or framework to help you, uh, clean or transform your data?

Data-metadata-issue-fix

Interviewee

Um, not, I think the only way that I know is the, all the things that I do is by myself. In a very instinctive way. I say, Hey, maybe I can create a new feature, or maybe I can between these two data points, create an in the middle because this data has some kind of pattern, but not don't much libraries for that process.

Data-metadata-issue-fix

Interviewee

Well, for, for language model, I think there's the common libraries space or, um, LTK but, um, Yes. Or even in transformers, I even listen about something that helps you to generate new data.

Interviewee

Okay.

Interviewer 1

Thank you. Um, is there any other data quality issue we missed that you consider relevant?

Interviewee

Uh, yes. Bias bias of the society of any language model. They, you will find a lot of bias in, in the text. So there is a lot of sex seeds, uh, rashes, um, and adults that has an impact. Cause when you, you train a transformer, for example of language, uh, is, is easy to find things like a, a woman work as.

QM-fairness-issues   Quality-Measures (QM)

Interviewee

And it says like a made or news or things like that. And when you are fine tuning the model, you need to be very careful and avoid that your model, keep that kind of pattern. So the bias and, and the thing is when you try to generate model often, you don't know which that contains you model entirely. So if you don't and take care about these.

QM-fairness-issues   Quality-Measures (QM)

Interviewee

If you don't take care about this, it can make your model perform well bad with this bias of the society. So the quality is not only about have good data is about have data that has some ethical sense. So that can be a quality issue. I, I remember a transformer that. We fine tune and we test it with these kind of processes and then the, well, the fine tuning was working well.

QM-fairness-issues   Quality-Measures (QM)

Interviewee

Cause we, we start to see that the model lost the bias with the time. Okay.

Interviewer 1

Super great. And, um, how did you ensure that the, the bias was removed

Interviewee

from your model by testing cases that we consider? Like we know that the model has the bias because we're training with Wikipedia and the internet and internet is made mainly for.

Interviewee

And white man, the things. So how you say in this case, they have dots on model cards that allows you and to, to know which data was to train the model and when, um, all the model is performing and it makes like, um, some recommendations, but once you train your model and you taste like, uh, to, to, and you taste to, to look for these patterns or bias, but you can see that it's not.

Interviewer 1

Yeah, sorry. I didn't said the, the right sentence. What I meant is, um, how do you remove bias from your model or how do you remove bias from your data set?

QM-fairness-issues-fix

Interviewee

Okay. In general. Yeah.

QM-fairness-issues-fix

Interviewer 1

Well, for your case, for your transformer.

QM-fairness-issues-fix

Interviewee

Well,

QM-fairness-issues-fix

Interviewee

so the bias that I mentioning is a bias that was initially on the weights of the transformer, right? Okay. I see the transformer comes with that bias by default. So much trained with a bias service. And what I'm interested is in the weights of the transformer, uh, in all case when I was working low and I started to, to fine tune these transformers, I just reset the weights from zero and I started to use it with code.

QM-fairness-issues-fix

Interviewee

So in that way, I don't want to have any, uh, weight related with, um, human behavior. I would. Ideally my, my objective when I was working on blur was to create, um, tools that were with languages. So how to complete, or to, um, say a tool that created doc strings by reading the function that used write or semantic search by

coding.

Interviewee

So for example, a function that generates prime numbers, which generates a function things.

Interviewer 1

Okay, thank you. Um, how do you evaluate the quality of models? And as a reminder, quality is not only defined by ML performance, but also by other aspects such as explainability fair, scalability.

Interviewee

You name it. Yes. Well, it's the first thing that you have in.

Interviewee

Um, the usual metrics that you use for a semi machine learning model, but at the end, the, the way that you will evaluate if you project is working is with the, with the, um, some increase in the budget or some, uh, metric that does the, the end metric at the end of the process. So if you are able to generate money for a company, then your project is a.

Interviewee

I will make an example, uh, a company in which I used to work. Uh, we wanted to make a recommender system, of course, for the students. Ideally the students will follow a path, but maybe by looking the path, I can find some patterns and compare with other students and recommend some new courses to, to these, to, to keep, uh, better engagement on the.

Interviewee

So the first thing that you need to ask is what is the impact of the recommender? And here you don't need that model. You can make a random recommender system and with test it with maybe testing and it worked. The engagement of the platform was, uh, uh, was increased. Cause the people say, Hey, platinum now counts with that recommended assistant that's created is recommended in the new things.

Interviewee

We don't have any model and. So that that's success and, and the metric was the engagement of the people. So you, you need to think in metrics that, as you said now, not related with the model, but you need to think who is exactly the stakeholder and why is the objective? And then once we start their recommender, very simple recommender basical on similarity, uh, while the people continue to be happy, the increasing of the engagement was not much different.

Interviewee

So just the feature, it calls an impact. To have the recommended working well, just allow to students to continue to, to play in the platform. Um, so here the metrics were not important at all at the beginning was more the concept, the idea that you want. Um, often I, today I was talking, uh, about the project and I say like, I would like to have a way to include the.

Interviewee

So for example, this, in this case, I want a better engagement in the loss, on the function of the model, because that's the objective of the, I wanted to do a classification model, but instead of use the useful, um, loss function, um, maybe something that kept the business metrics on that can be created. Uh, but I dunno, maybe then some.

ME-metrics-a-are-poor-proxy-of-quality

Interviewee

People that is working in something like that. But, um, so yes, you need to consider several metrics at the end. I will say business metrics or, or even metrics that evaluate the performance of, of your model. Maybe your idea is too complex and it will take too many months. And the monitor you will, uh, waste developing the project.

ME-metrics-disconnected-from-user-adoption-fix

Interviewee

Doesn't work the, the income that you will receive. So at the end, most variable metric. Income or something related with a, um, generated value.

ME-metrics-disconnected-from-user-adoption-fix

Interviewer 1

Okay. See, thanks. Um, so you mentioned KPIs and, and you said that you, you had the idea of using the KPIs, uh, to, to, um, train your model. Yeah. Did, did you succeed or.

Interviewer 1

You did?

Interviewee

No, no, no. It was an idea that it has today. So I want to see maybe there is a way for letter I could, I can use in the loss function of model. I don't know. It's just an idea. I don't know how to do it.

Interviewer 1

Yeah, it looks, uh, it looks great. Yeah. If it works, tell me

Interviewee

for, to research about it and study your invite.

Interviewee

It would be great to part. All right.

Interviewer 1

Um, so moving on to the next question, have you ever used a existing benchmark model for quality aspect to evaluate your model? So benchmark model, have you ever used some, some,

Interviewee

the first thing that comes in my mind was, uh, the logistic or being Aion. I think those are very good benchmark to start any project.

Q4.2

Interviewee

Um, Nowadays, uh, there is a lot of libraries that are very automatic when you are making models. So for example, by priorities are auto ML library, and it's very straightforward database, you and place some functions. And then it takes for you a lot of, um, models like L G B and decisions we run four is, and you can

quickly achieve good performance.

Q4.2

Interviewee

And those for me are very good, uh, benchmarks

for

Interviewee

any purpose, then you can start to do more complex fees.

Interviewer 1

I see. I see. What's the name of the library? Yeah. I mean by how do you

Interviewee

spell it? Yeah, I writes develop it some Canadian faith. Okay. I see.

Interviewer 1

Thanks. All right. Um, have you ever assessed a quality of ML model prediction with the user of your, of your

system?

Interviewee

Uh, can you repeat the question again? Good.

Interviewer 1

Have you, oh, sorry. Have you ever used, uh, sorry. Have you ever assessed the quality of ML? um, sorry, ML

model, model. Sorry. , I'm a bit tired anyway, um, with the users of your system. So you check if your model

was behaving correctly with the users, with its users,

Interviewee

um, in which case.

Interviewee

Well in , uh, we create, uh, some char model to predict when the char immediately, um, the, the S wanted to use for, for, uh, increased engagement of the people. But the thing here was that I, I asked it to then to, to please split the population into, uh, into, in order to do AP testing. And now if the model was performing well or not, they apply in the entire database.

Interviewee

So that is, uh, often that happens, uh, Because they, they want money, but at the same time I wanted to taste it pay model and all my hypothesis will work or not. So what, at that time, for example, I wanted to do it. I didn't have the opportunity with the recommender. And the only feedback that we start to have was like in Twitter or social network, like when the people were telling like, Hey, can you recommend there is doing very well.

Interviewee

So that's like the. Um, I left the company, so I never knew what, I never just knew what happened with the project itself, but I, I think there's very simple ways to test if your model is performing well or not. Um, right now a project in movie, we are testing with the sales. The results of our model, but the, the validation can takes like a two, three months until we start to have very good re results.

Interviewee

So I will say my experience, I deploy models, but, uh, the testing part is being hard. Uh, now in the new project movie, I wear, basically the thing that I will do is test is analyze the, the results of the model. So incredible that seven years, this will be the first time that I. Start to work only in the, what can I say that day after deployment process of the model?

Interviewee

Yes, the monitoring phase. Yes, because I have participated lot by designing or creating models. Even my position in was so far designer and designed so far and that still have to continue the process. Okay.

Interviewer 1

Let's see. Um, have you encountered any other quality issues, uh, during the evaluation of your model

Interviewee

during the evaluation?

Interviewee

We have other questions. So, no, I, I would, maybe the only thing that comes in my mind is the interpretation of the, of the outputs of the model. Um, It's very hard to, we're gonna say that vulgarize the, the concepts to the business and to the stakeholders. Normally from the data science point of view, it's easy to understand what you are doing, but, um, its not that quality issue, but its hard sometimes to explain to the, your ideas and if you're, if you.

| QM-explainability-issues | | Quality-Measures (QM) |

Interviewee

You don't understand very well the business, because that happens sometimes. Um, you can make that it can happens that all you work or model doesn't work for the purpose of something. I have been involved in cases where you do a lot of job. And when you start to see if that was the need of the requirement, uh, we were far of the real object.

Interviewee

So understand the, the mean the, the problem's not easy sometimes from data science. So from the business perspective

Interviewer 1

of question around the team of, uh, learning software system deployment about four of them, uh, so how and where are your models?

Interviewee

Um, how

Interviewee

well, the, the model that is being deployed by, for example, in the, the, with move, um, is in, um, is in databases. And we are using ML though. Basical. Okay tools. Okay. That one that is using ke um, this previous model that I have done, you said also I something, well, the pipeline we were using Jenkins, we were in back in, uh, some, well, basically.

Interviewee

Okay,

Interviewer 1

thank you. And, uh, what are, sorry,

Interviewee

it's some, some other projects where I use it, uh, the Amazon work services. Okay.

Interviewer 1

Yeah. Perfect. Um, what are the challenges that you have encountered during the deployment of machine learning software system?

Interviewee

Um, the like a, a very time that we need to retrain the model and you create a new version.

Interviewee

Uh, I, sometimes you don't, I, uh, you want to test this. If the new version will perform as well as they author one, but because you need to do this like, uh, every month, or you need to do this sub, there is the risk that you model under performance and guru part of clear vision. So I was using tools like evident.

Interviewee

To, to manage, uh, see differences between me, my, my data sets or models.

Interviewer 1

Okay, perfect. Um, did you ever have a model that performed well locally, but once deployed poorly?

Interviewee

Um,

Interviewee

Well, yes, I cannot say that it, when she's deployed work poorly, but. When you do the validation you expect to have, I don't know, 90% of performance. And then when you are measuring the perform, the output is 80 or less. That's something that is happening with the project of Company 3. Um, there is one of the, the plan that we have move yet that one with kero and the churn model that I designed for for PLA C and here, the problem is because the way that you.

Interviewee

The, the model and the metrics that you are using maybe are not the same as the met as the business metrics. So one thing is create a laboratory when you play with all the things, but once you are testing in the bus from the, in the reality, um, you mentioned it started to change. So that's why I mentioned that I would like to have a lost function that includes a business perspective.

| ME-metrics-a-are-poor-proxy-of-quality | | Model-Evaluation-issues [ME] |

Interviewer 1

Yes. Okay. I see. Thank you. Um, have you encountered any other quality issue with your model or system during the deployment phase?

Interviewee

No. Okay.

Interviewee

All right.

Interviewer 1

Um, so now we are in maintenance. How do you ensure that the quality of a machine learning software systems stay the same, uh, through time?

Interviewee

Uh, well, there is some testing that you can do. I don't remember exactly the name of the testing, but to compare that your model is producing the same distributions, uh, your data, it doesn't having too much changes due in the.

Interviewee
Uh, again, data shift, something like that. I don't remember exactly the concept. Uh, so this, this idea, so keep track of their data on their output of the models. Okay.

Interviewer 1
Um, have you ever encountered issue with data during the maintenance system?

Interviewee
Um, well in, yes, for example, the. The one project with Company 3, the COVID, uh, um, period impacts the performance or, or the behavior of dataset.

Interviewee
So you need to do predictions over one year, but to predict on this year, I will use the previous year that was impacted by the COVID. So the data is not behaving us in the previous two, three years. So that case, that case is very interesting because of that. And the thing is you cannot correct. There is some kind of tweak in the, in the variables that is dis suspected that

Interviewer 1
yeah.

Interviewer 1
And how do you these issues and how can you prevent them from opening?

Interviewee
Um, so you can reweight the project, or just be aware that. Model is using this like a bias, uh, data set or you can yes. Like, uh, rebalance your data set. It depends on the cases. In, in our case we have a lot of data. So at then I did some like validations, I trained.

Interviewee

The same model with data until I don't know, 2018. And then I joined with data until 2020 and then until 2022, and I started to compare the AUC or any matrix. And the behavior was like the same. So for me was not a issue, but this is because we have too much data. So the, the model was regularize it because the amount of data.

Interviewee

So that, that was very good. Uh, but I don't think that can be the case of other projects. So then you need to rebalance the data or wait different the, the, the months that you are using, but you need to find a strategy or shift your objectives can be a, so. Yeah.

Interviewer 1

Perfect. Thank you. Um, so did, okay. So now we are finished.

Interviewer 1

We, we cover pretty much all the phases of ML pipeline. Now I have a, like two or three question left for you. Um, so did you ever add issue with one of the following quality aspect, fairness, robustness, explainability, scalability, or privacy?

Interviewee

Uh, all of them, I will look. The first one was you, you mentioned five.

Interviewee

Okay. Uh, yeah, five the first

Interviewer 1

one, sorry. Uh, we can go over each one. Uh, fairness. So you mentioned it

Interviewee

earlier writing. Okay. Well, fairness from which point of view

Interviewer 1

you have a model that, um, it, oh, sorry. I said, okay. Okay. You have a model that, um, As different group of people can, it can be different sex, different race, different ethnicity, anything, and your model must behave fairly toward each of these person. So for example, a bias model could say, like you said earlier, um, a woman is more likely to be made or I don't

Interviewee

knows.

Interviewee

Yeah. So, well, I, I didn't have that, uh, cases in general in this. Because, um, my, my models were not applied given in these, uh, I know characteristics of the people. Um, but I used models that were biased by this kind of, uh, Evans. And I assume you the best way just to like, reset everything that the start from zero, if you have enough data to, to create a new model that's so that's preparedness.

Interviewee

Um, maybe the thing here. Try to rebalance your, your data and select what make the data that represents what you want to do at the end. Very hard. Depends on the case on the project. So that's for fairness. And you mentioned author robustness. Explainability. Yeah. So as a, a very good thing to, to mention, um, for the pro project I design.

Interviewee

Very nice, like a framework, uh, which have values. And ideally, uh, the idea was to, to don't only deliver a score to the people was also to explain to them why the model is behaving like that, or what are the features that are impacting the, the, the, the output. And that was great because the people that was using.

Interviewee

The output of the model was in some way, understanding the conditions of the prediction. And we calculate shared values by all the prediction of the model. So that's very, very nice to have in, in general, even for data science.

Interviewer 1

Okay. And did you ever use models that were not explainable and this was an issue for you?

Interviewee

Well, it depends cause. Sometimes you need to deliver a model. You do, and you forget the data you continue in other area. But, uh, I will say that always is important to, to keep track about why the model is behaved in that way. Um, well now I remember one model, the idea was to classify large documents. So we did a classification model, but.

Interviewee

It was missing a part about why the model tech does decisions. It's very hard to understand what patterns the model is find in the text to, to perform a prediction. So it becomes a very black box and that makes that you cannot explain to, to the business. And you don't know exactly if you can trust the model or not, you will trust because it's producing very good results and the, the metrics are good, but.

Interviewee

Does it work for all other cases? It's, it's very hard to know. So yes, open that black box for me is very important. Mm-hmm

Interviewer 1

perfect. And did you have any experience when, with one of the other aspects, uh, robustness, scalability and privacy,

Interviewee

uh, it's liability. Well, uh, most of my projects, uh, leverage, uh, high scalability.

Interviewee

Most of them are experiments. And as a data scientist, I, I, I do a lot of hypothesis, but no one finishing something very, um, At introduce the PR project that is becoming very well practice, I won't can explain is the idea is to credit maintenance events for a large firm that produce engines. So the engines need to, uh, make,

we need to, the plane needs to do some maintenances after some number of like hours for these, uh, airplane engines.

Interviewee

And we are doing the prediction about when these Evans will occur and. Um, they are working with McKenzie. There is field and they want that all the pilots are very robust and right now they, they, the pipeline that we created super well decided and hates it and is able to be already and deployed. And that's, I will say one of the few cases when I arrived to this point, uh, about Rob robustness.

Interviewee

Well, sometimes because I work in a lot in startups environment, you don't require, uh, a model that perform 100%. If you have a 80% of performing well is acceptable for the first ideas. And sometimes the companies just forge that and it start project. Okay. I would say that's hard to, to achieve the point of what you are able to scale a project and keep maintenance over it.

Interviewee

Even, even for moving. I, I believe that us as a consulting firm, we are doing a really good job by deploying or creating, uh, draft models versions. But then I, I feel that the models are used, um, Forget it with a client and we continue with another client. And as a consulting where one of the first aspects is in order to have more incomes, you need to make a continuous development of the things.

Interviewee

And I believe that that's a thing that even in these companies missing, we are missing the bring continuity to the projects that they are, that are being deploy.

Interviewer 1

I see. Yeah, that's a good point. All right. Uh, so I have two questions left for you. Uh, in your opinion, what is the most pressing quality issue researchers should try to solve?

Interviewee

What is the most, uh, important quality issue researchers or should try to solve? Yeah, sure. Sorry. Should try to solve solve. Oh, okay. Research should try to solve. In the war in the entire pipeline,

Interviewer 1

uh, to solve quality issues and learning of our system. So yeah, in the

Interviewee

entire pipeline machine to the fuel,

Interviewer 1

any issue you, you have mentioned so far, basically,

Interviewee

You know, for me, the first issue that, um, and it's a fragment question that I do every spring in any development. Is this I'm, is this solution, uh, solving the needs of, of the really need of the project? I, I do this question all the. It's something it's very funny because you are working all the weeks, all day long, but at the same time, I, I, I hesitate all the time.

Data-no-feasibility-analysis

Interviewee

If this will bring value or, or not, as I say, it's science. So you are basing your basic, your solution in hypothesis that you never know how it'll perform. And. That's a very funny thing. You dunno if your solution will work on those slope, keep in asking me what, what are, um, if I have a really good understanding, understanding of the business perspective, um, I see often that you can create models or, or try to model their reality.

Data-no-feasibility-analysis

Interviewee

But sometimes there is a unmatch between the solution and the reality. I, and that's maybe why some models are not getting deployed at the end is because maybe the understanding of the reality of the real problem is not being very well classed by the, by the model itself. I know that there is a lot of parts of.

Data-no-feasibility-analysis

Interviewee

So that you can improve as the data. So you can make a lot of improvement of the data, but at the end that doesn't define if the solution will be achieved or, or not. So from my point of view, the thing which I deal every day is the business understanding to have a very good business understanding and how to interpret that and transform in a.

Interviewee

In, in, in math and code and statistics and make that work and work, present the reality and have an key that I can in which I'm simulating the, the scenarios. And to know if the output of all of that will be the real value that I'm looking for because generate bank for the companies is it's more important.

Interviewee

And sometimes I Don. I, I feel that it's easy to lose the, the track that you are doing a lot of things. So I know that is quality issue, but my main everyday problem, because quality at end can be the, is the reason of other things. So if you, those is not performing well, it's because your data is not being well.

Interviewee

Uh, process it, cause your model is too complex or because you didn't understood the business and maybe you go with a solution that is not a key value of what their, their objectives or the KPIs are

Interviewee

trying to go to the roots of the problem. More, more than seeing about the quality itself. I see.

Interviewer 1

Well, that, that's super interesting. Thank you. Yeah. Uh, so we're done with the, uh, with the interview and that's all for us, so thank you a lot, Interviewee, to, to take part of, uh, our project

Interviewee

will help to this project.

Interviewer 1

I, I think so. I think so.

Interviewee

Yes. Let you know. And well, let me know if anything I can help. I'm really sorry to, to be late.

Interviewer 1

I it's it. It's it's okay. No worries. I'm very happy

Interviewee

to participate in this and I thank you too much to, for this entire hour. Oh no, we thank you. And keeping Interviewee, I would like to see that the whistles of all these interviews that you are doing, maybe there's something.

Interviewer 1

Absolutely. All right. Thank you, Interviewee. Have a good. Thank you. Thank you. You have a good

Interviewee

day. Thank you.

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