P-4

Depends on the application again I will say, like more than the application I will say the domain field, because if we are not familiar with the domain field, you might not be aware what is more important for the customer. Perhaps according to your experience you might be favoring some particular requirement that is not as important for you.

P-4

Sometimes it’s a lack of a documents that exactly present which are these non-functional requirements or the requirement of machine learning system should compatible.

P-2

Yes, challenges that machine learning will not behave in the same way. So, I don’t know how you wanna measure that. If you wanna like to testing on run time so the time or if you wanna keep a complete lock out of how the system behaves to understand their problems. This is really tricky, I think. Because usually implementation will behave the same in same situation. Where as machine learning could behave differently depend on how it trained and how it perceives, how it interpret sensor information as well, all this aspect make it really difficult.

P-2

You can test the system when it kind of done then do you wanna test for each scenerio. So the context operational design domain is important but even then so many things that play into this. Sensor Interpretation, how it’s implemented how the relationship is between different component like they might be again a problem or communication problem maybe that again if you don’t test for these things.

P-9

Measure? Not that I’m aware of. I mean that, testing the system of course and based on those tests we decide whether it’s safe or not, but in Machine Learning I’m not aware of any possible measure yet

P-9

I mean for safety requirements; it will be too much the way I mentioned for hardware. If you have the strictest hardware requirements, I think is that you can only have a random failure as a probability of 10^-9 per hour. So, it’s extremely extremely low. And I think something similar you can define for neural networks probably as well, that the only problem then is of course that you have to provide a neural network with plausible input data that represents the operation that you actually want to do with the neural network. And that is what I meant when you need a lot of testing which is expensive.

P-9

Yeah, as I mentioned before, the biggest challenge I think is to provide a correct and plausible data that would be expected to also be representative of the used cases where you will deploy your system with the Machine Learning model. If we only have test data that is academic or does not represent the actual operation of the vehicle, then of course you cannot guarantee safety, because we couldn’t test the neural network properly. Plus, you must be absolutely sure your test data and your training data that you used to are not in any way coming from the same data set. Because then you a very quickly run into the risk of of over fitting your training data if you don’t have independent test data then you do not realize that you did over fitting and at will be two significant funny results from your neural network that are potentially extremely unsafe.

P-8

First of all, be aware that I measured the what the right ones, because you know how it goes. You start measuring accuracy and everything is good and you’re happy, and then one of the customers is not satisfied and they complain and then say, let’s measure accuracy per customers. And then times go by, and the system crashes. OK, we should take care of that. And if it crashes that you say that MS Word is a Malware. Sometimes it feels like everybody has a list of accidents, they try to avoid them in future.

P-10

Challenge is inherent in many cases. First challenge if the service is actually affecting a very vast majority of people. And trying to understand the requirements, it’s a random process, try to understand the requirements in a meaningful clustered way, like market segmentation. So, it’s very challenging also that how we will define all these criteria. It’s it becomes a statistical problem. So, it is statistical problem and majority of data, the randomness inside the data, the randomness of the stochastic procedure of this whole scenario of certain software makes it complex that how we will adopt the metric.

P-10

Some metrics will work for some system component of the system or the segment of customers or some of the metrics will not give much valuable insights of the system for that certain part. So, we have to adopt several different metrics to understand the whole picture. But it’s not always easy. Because I think there are many randomness and it’s not always structured like the data and the processes are not structured and the process itself is random.

P-3

I guess getting the benchmark of the measurement, the baseline they should be measured. For example, is this model accurate or not compare to what. Often these baselines are not properly defined. That is one challenge.

P-5

If you compare the functional requirements we are probably way behind the when it comes to non-functional requirements. We don’t have the same strict system for that as we do in the functional requirements. Which is perhaps something we should get back to that.

P-5

Since I work mostly with the machine learning, I would say we can setup the rules for how fast something should react. And we do have some requirement around that we couldn’t take more than this and that. So, we do have some of that written down that but when it comes to the others things that I talked about, it’s mostly been around discussion. So, we all aware about this problem but it’s not that well documented. It’s not really part of our development process I would say.

P-1

yes. You have some statistical bias that could be quite big in some cases. If you work with historical data and you create machine learning models then you might be quite bias. So the challenge is that you need to do some prospective study to actually verify the usar can be benefited by using these ML system.

P-6

The main challenge is to find an effective way to measures it.