

Looming Threats - transcript of presentation video

Nick: Team 4 Looming Threats, please give them a hand.

Mark: Good afternoon, Simon and I have the absolute pleasure to be presenting on behalf of Team Looming Threats.

So, one of the, I guess key trends that has been happening in the past few years is, we've seen numerous case studies of financial crime, money laundering that have really demonstrated the importance of sharing good quality intelligence, and coming together and collaborating to really effectively disrupt financial crime. This has really led to in the past few years, an increasing number of public private partnerships, such as JMLIT here in the UK. And if you kind of look at what these public/private partnerships are doing, ultimately at the base they are coming together to collaborate to develop, and share good quality typologies and we've really seen in the past few years that this is having results.

We've seen for example, in the UK, the latest economic crime plan really referencing the importance of public/private partnerships, but being realistic for a second, and thinking about the scale of financial crime, the size and the scale of the existing public/private partnership remains quite small and it's still quite manual.

So, our team, has really kind of seen an opportunity this week, to introduce technology, to really kind of bring scale and accelerate the success of these public/private partnerships by introducing innovative technology and the specific technology that we are going to be talking to you today about is something called **federated learning**, so we will only see federated learning being able to bring the scale to, in the first instance, just sharing and developing typologies, but then with that capability being able to really kind of facilitate and accelerate the increasing number of participants joining public/private partnerships. And ultimately starting to increase the coverage so leading to better outcomes and I will explain what outcome is for the industry as a whole.

So, I mentioned this idea of federated learning, probably a number of you are kind of asking what that is. So I guess I will start with a bit of technical definition. Federated learning enables multiple institutions to collaboratively train a model without sharing their local data. Again, probably a number of you are saying, well what does that really mean?

So, I will kind of step out of financial crime, and model language for a little bit and just give you an example in the real world. So federated learning for example, is being applied in the healthcare space, it's been used to detect things like cancer. So, if you think about hospitals, they have images, they are using those to try and detect cancer but they can't kind of share and collaborate amongst themselves. So, federated learning really provides a way to build a global model, share insights, good kind of ways to detect cancer without actually sharing the sensitive data such as personal images.

So, stepping back into the AML space for a second, and thinking about how federated learning could be applied in this space. Today, each institution is building their own detection models such as transaction monitoring system, and federated learning could really provide the kind of rails or the mechanism to actually share the kind of insights and best practice anyone institution is building up.

So, what have we been doing for the past weeks, so I guess the starting point for our team is the synthetic data that was provided, and that synthetic data contained customer information, account information, transactions for a number of different institutions, and like any bank in the real world,

the starting point is building a detection model to try and identify interesting patterns that might be indicative of things like money laundering. So, this is exactly what the team did, they treated it a bit like a real-world problem, they explored different types of machine learning modules, graph techniques and worked with the data provided to kind of build and test and validate a model for a single institution's synthetic data. But really what we kind of looking to do then is to show how we can share the kind of capabilities and the insights that have been built through federated learning. So, some of these are the other institution data and the synthetic data sets.

So, Simon, at this point, I think I'm going to hand over to you, to run through the demo, to kind of bring this to life a little bit more.

Simon: Thanks Mark, so the scenario we want to show involves two banks, one bank will have a fairly well-trained model, and the scenario we want to show is that, there will be a second bank that has a slightly less well-trained model, maybe it's a new installation, that's had few iterations of the model.

So, Mark, so we are looking here, so the first bank is Home Counties Bank, a fairly typical case management system, a list of alerts, the top alerts is got a structuring typology, it's got a fairly high score of 0.98, so we can go into the alert, we can see some of the statistic that an AML analyst might typically look at, for example money in versus money out, that sort of thing, the first pattern we see from our advanced graph analytics is structuring. We are looking at high number of income transactions, like 15 here, low average balance over the period we are looking at £1400, we can actually visualize that, with a sub-graph, in the middle here we've got the account, you can see, it's a sort of a hub shape, it's fairly easy to see, by eye that something is going on there. If you go back to the alert, also you can see even in the world trained bank we can see that SAR confidence goes from 0.86 to 0.98 even with the world trained module.

If you now go to the second bank, which is called the Berry Western Bank, if you look at the third alert here, it's from company called the Beasley Group, it's also a structuring typology but for some reason their score confidence is only 32%, so we are going to kick off around the federated learning, what's happening here is that banks models are actually exchanging model configuration between them.

And now that's completed with back of the Berry Western Bank and we can now see the Beasley Group Company is the top alert and the score has gone up to 96%. if you look at the actual alert again, we see the structuring pattern, we see a high number of transactions, we can see the average balance is low over that period, but what the important thing we can see is that our score is gone from 32%, we've done the collaborative learning, we've done the federated learning, we haven't traded any transaction information at all, and we've gone up to a score of 96%, which is where we think it should more likely it should be. Thanks - Mark.

Mark: Fantastic, thanks Simon, so hopefully that give a bit of a flavour of the power and potential federated learning. So, taking a step back, and we've just kind of shown you how this might work, between two banks. Just think about what might mean in a kind of the bigger picture here. So today we have again multiple financial institutions building their own model, reporting SARs, and the feedback loop, through the FIU back to the individual bank is quite slow, it's through documents. So really what we are introducing is federated learning kind of protocol, is something to bridge that gap, something to accelerate the feedback loops, and provide the kind of mechanisms to share some of this typology information.

So, I guess, if we were kind to lists 5 things that we really want everyone in the room to kind of take away, I guess the first one is that, it's important to stress that with federated learning, no personal data has been shared here, we are sharing the insights, the typologies, not the underlying local data. This in the grand scale of things compared to things like big data warehouse, lakes, it's much easier to implement, this is about creating the channel for each bank, kind of speak to each other and share those kinds of key insight around typologies. Anyone can participate, so we kind of see an open standard, where everyone can communicate through the same protocols, and really start to benefit all participants.

So, any one bank can gain the insights from anyone else that is participating with this network, so we kind of seen that its really helping to improve the overall quality of detection.

So, we've got 4 seconds left, so thank you very much to the team, it's been absolutely fantastic.