ADGMA RESEARCH CENTRE ARTICLE SERIES

FEDERATED MACHINE LEARNING AND ITS APPLICATION IN THE FINANCIAL INDUSTRY

'Potentially applicable to all distributed data sets protected by privacy laws.'

Building on the recent publication of the white paper, **Leveraging Artificial Intelligence to Enhance the SME Ecosystem in the UAE,** Peter Ware, Head of Research & Development at the ADGM Academy Research Centre, sat down with Professor Duan, Head of the Asian Institute of Digital Finance, at the National University of Singapore and Richard Hills, Associate Managing Director at K2 Integrity, based in Abu Dhabi, to explore potential applications of federated learning in the financial industry while taking a broader view on the evolution of technology and some of the barriers to adoption of innovative solutions.

Kicking off the discussion, Professor Duan described the approach taken for all research undertaken at AIDF – first identify the issue to address, rather than starting with a specific technology solution. For the SME research, federated learning was seen as the most appropriate solution to meet the requirement to build advanced analytics using data across different sites that were subject to privacy protection.

For those who like definitions, federated learning aims at training a machine learning algorithm, for instance, deep neural networks on multiple local datasets contained in local nodes without explicitly exchanging data samples. The general principle consists of training models on local data samples and sharing parameters (e.g., the weights and biases of a deep neural network) between different local nodes to generate a global model shared by all nodes, as well as produce enhanced local models catering to varying needs.

RELEVANCE TO INDUSTRY

Drawing from extensive industry experience, Richard Hills picked up on the concept of building advanced analytics using distributed data sets subject to privacy protection and reflected on the challenges facing the financial industry. He highlighted that when discussing credit risk, fraud risk, and the range of risks arising from financial crime, it soon becomes clear that these risks are not only being dealt with in silos within a single organization, but also across the different organizations, possibly

operating in different jurisdictions, that make up the global financial sector ecosystem. All participants face the essential issue of insufficient data and consequently the inability to act proactively.

If we think about the ongoing digital transformation of the financial sector and the disruption that Fintechs and technology have brought about, banks have had to react by doing things in a smarter way and by improving their efficiency. Yet one of the effects of digital disruption, apart from impacting customer experience, is that it's also starting to impact growth whereby the need for digital solutions and the development costs incurred are coupled with a downward squeeze on profit margins. From a digital transformation perspective, there is a huge benefit to be gained from the power that new technology can bring to solve existing problems.

TECHNOLOGY, SMARTER AUTOMATION, AND THE DIFFICULTIES WITH DATA

Echoing Richard's thinking, Professor Duan agreed that sometimes we get so excited about technology for technology's sake. There are a lot of traditional issues that need to be addressed and as firms embark on a digital transformation journey, the issues they aim to address essentially arise in a traditional space. Traditional finance is already largely digital, just not necessarily efficient, and often encumbered by legacy technology. So, another way to describe this requirement is encapsulated in the term 'smarter automation', because what needs to be solved typically are smarter automation projects. We can't turn back the clock on technology, as technology is all pervasive. So as technology evolves, the solution also rests with technology.

Financial crime also falls into this category, and the need for federated learning goes back to the fundamental issue that data cannot be shared directly due to privacy preserving laws and for competitive reasons (from the banks' perspective). In practice, there is one very important feature of federated learning that we always must confront, that being the fragmented nature of data. Fragmented due to organizational differences or preferences in the developmental stage, or simply by choices made to collect different kinds of data and different data practices. Consequently, it becomes almost impossible to harmonise data, and even more so to get involved parties to agree on how to harmonise.

When we deal with a new style of data analytics to solve smarter automation, we must be mindful of the fragmentation issue. Even when centralized, data tends to be fragmented in many ways. Technology can, and should, address that. That's why in the SME white paper, we addressed what we refer to in the credit space as alternative data.

Credit has a long history, and there is some consensus on the leveraging of relevant financial metrics. However, there are many things we don't agree on across organisations and amongst individuals. At least within a single organisation, when there is no consensus, decisions can be mandated. However, across different organisations this is almost impossible. So, technology, whether federated learning or others, must address this fragmented data reality. Coming back to the technology used for SME credit risk, there was a need for some fine tuning to deal with those attributes with which we could all agree and those attributes with which we could not, but this type of grand analytics, when applied, precisely reflects the diversity across one or multiple institutions.

From a financial sector standpoint in the UAE, Richard fully agreed and noted that taking any three or four banks in the region, each one of them would be at a different level of digital transformation. Some may be completely analogue, based on their business model, products and customer base while others may actually be 'digital masters'.

From a regulatory perspective, the ideal would be to get everybody up to a certain level of digitalization. However, the steps that an organisation needs to go through to become a digital master requires a significant effort. The transition from analogue through to digital ultimately enables the benefits of the digital transformation process, and the application of methods and technologies such as advanced analytics, machine learning, and federated learning.

The digital transformation journey is a complex process and not all financial institutions have the capability, funding, or the appetite to go through it. But, if we consider all the benefits that a federated machine learning model in the financial crime compliance space would bring to the industry, it would, at a minimum, enable a certain standard for all participants. The prospect (of such a machine learning model) may incentivise the more conservative players to catch up, even though they might not believe that they need to. It could also help to bridge the gap with the pioneers, bringing equal benefit to all participants.

A PERMANENT STATE OF FLUX

Professor Duan observed that, in a sense, whether pioneer or laggard, all organisations are constantly playing catch up on the digital transformation journey, and it makes sense for every party to advance. However, not all organisations will advance to the same level. Considering this, for technology to be truly useful it must be able to take account of and accommodate this constant state of flux (and evolution), whereby some degree of harmonization will always need to happen.

If we consider the use of federated learning to address an issue such as financial crime, yes, it's possible, but it will always be a work in progress and for Professor Duan, therein lies the beauty. The beauty of always having something to improve. Technology is constantly evolving and what was not previously possible, suddenly becomes possible, which makes it all very exciting.

Financial crime is clearly an area that could experience real benefits, particularly given the fact that institutions can be mandated by the relevant authorities to co-operate to a certain point. The benefits of collaboration are real, however, in general there is often resistance in taking the first step. This is something that regulators could initiate while at the same time applying privacy preserving measures. Expectations would need to be clear given the reality of the fragmented nature of data, however, the results would be far better than without such an initiative particularly given that the issues relating to financial crime are well defined and articulated.

Building on this premise, Richard added that it's one thing to monetize and use data across different institutions, but we shouldn't lose sight of the value of utilizing different data sources that are currently being dealt with in silos in individual organizations. In terms of federated machine learning and taking a very simplistic approach, it would be a shame to have separate models for credit, fraud, money laundering and sanctions, as these are all areas of compliance and/or geopolitical risk that essentially inform and feed into each other. If you could bring it all together, the result would be a far better and richer view.

FACTORING IN THE HUMAN ELEMENT

In response Professor Duan highlighted that the usage of a model depends on the purpose, and while there are definite synergies and interplays across credit and financial crime, realistically, we can only take it one step at a time. Relating these two separate functions to an organisation, they are typically performed by different departments, and convincing two departments to collaborate is not always easy. The point being that, even with the advances in technology, we still need to be mindful of people's behaviour and the human element.

In Richard's opinion this is particularly evident in the boardroom, where human decision making is not always data-driven but can incorporate other elements of experience, competition, and power-play. Be that as it may, machine-based analytical evaluation of data, providing a solid decision support mechanism, would at least remove the ambiguities in the interpretation of data and promote data-driven decision making.

From a technical perspective, this emphasizes the need for and the paramount importance of explainable, interpretable, and repeatable models. When we talk about models, we're not just talking algorithms; models are intrinsic to many forms of risk assessment. If we think of the model as a box, it requires input – essentially data – and provides output based on the rules and relationships between the data and the likely outcome. However, there will always be some degree of uncertainty in the outcome. If we introduce the term algorithm, we can see this simply as a means of execution of the model.

Focusing on federated machine learning, the concept of an explainable and repeatable model is no different and equally important. If the analytics performed on the data cannot be explained, then the technology or the solution is likely to trigger hostility and subsequently be more easily rejected. From a regulatory point of view, the decisions that are made have a concrete impact on customers and people, and therefore auditability and transparency is key. Customer protection must be top of mind throughout the development of such solutions. Practitioners, academics, and regulators all share this obligation given that ultimately, it's the customer that is funding these activities.

KEY TAKEAWAYS:

- Technology is all pervasive and perpetually evolving. Consequently, all institutions are playing catch up on their digital transformation journey (to varying degrees), therefore, for technology to enable truly useful solutions, we need to consider and accommodate for this constant state of change.
- Federated machine learning is potentially applicable to all distributed data sets that are protected by privacy laws, including financial crime detection. The challenges lie in the fragmented nature of the data, the obstacles around harmonisation and the industry appetite to collaborate.
- A key enabler towards successful collaboration both internally and externally to financial organisations is the use of explainable, repeatable, and transparent models that ultimately provide a solid and trusted decision-support mechanism.
- At minimum, the benefits that federated machine learning models could bring to the industry would enable a certain level playing field and provide equal benefit for all participants. In the face of hesitation there is potentially a role for regulators to play to encourage and support such collaborative initiatives.