Money Laundering Bulletin
Business intelligence | informa

www.moneylaunderingbulletin.com

FEATURE › TECHNOLOGY

Ever watchful – the AI & Machine Learning promise

Artificial Intelligence (AI) and Machine Learning will have a seismic impact on how money laundering is spotted and stopped, **Charles Delingpole** of ComplyAdvantage asserts, and explains how.

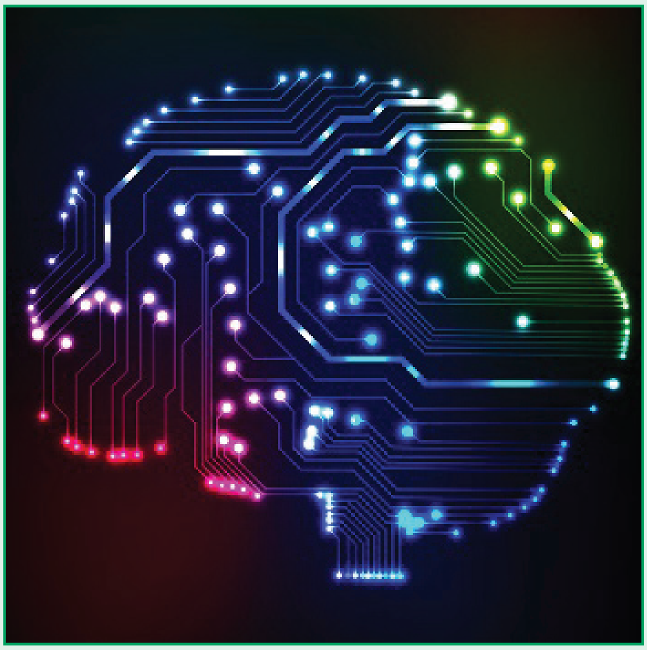
The current system is broken

Tackling money laundering has never been an easy feat for companies. There is a steady stream of news highlighting that despite the industry spending a huge amount of money on data and software to fight financial crime, there has been limited success to date in beating the perpetrators.

Furthermore, this challenge is growing harder by the day: regulations are becoming increasingly complex and vary across jurisdictions. External risk indicators are becoming harder to spot as the amount of information available grows exponentially and the speed of change gathers pace. Similarly, the volume of internal data to sift through is growing rapidly and evolving business models are creating new risks. This is further exacerbated by customer pressure to make decisions faster and all the while criminals are becoming increasingly sophisticated at finding ways to circumvent AML controls. As a result, it is unsurprising that money laundering remains relatively unrestrained.

Some of the key problems, and therein some of the most promising solutions, lie with the technology used by companies to underpin their AML programmes. Most of the systems used today originate from the 2000s and were never built for the scale of data or demand we see now. Systems built on infrastructure from even five years ago can quickly become overloaded. They struggle to spot anomalies and real risks within portfolios whilst generating huge amounts of noise in the form of 'false positives' with the associated cost overhead and negative impact on the customer journey.

In this article, we will examine some of the main elements of a risk-based AML programme and consider how AI and Machine Learning can be used to increase both efficiency and effectiveness:



- First, we will look at how these technologies can improve the quality of AML data that cover known risks from sanctions and watchlist lists, the political exposure of clients and adverse media;
- Secondly, we will look at how AI and machine learning can improve the ability to spot money laundering risks during the onboarding process, monitoring throughout the client lifecycle and when screening payments for AML risk;
- Finally, we will look at how these technologies can improve the monitoring of transactions for suspicious behaviour.

Whilst we will not cover all parts of a comprehensive AML programme, or provide an exhaustive list of all the challenges an organisation faces, we will aim to provide an indication of the wide range of applications these technologies can have in the fight against financial crime.

A 30-second primer on AI and Machine Learning

"Artificial Intelligence is generally used to describe the field of computer science where computers mimic the cognitive ability of humans, such as 'learning' and 'problem solving'. [1] In practice, this means that computers can perform human-esque behaviours and make predictions without being expressly programmed to do so. For example, you give the machine a goal, such as to win a game of chess, and it works out the moves to best achieve this.

"Machine learning is a subset of AI which involves training computers over time to make predictions and judgements based on the information and data they are fed. Simple regression models use two variables to predict likely outcomes making them prone to error. For example, looking solely at historic house price data - a house that has three bedrooms will then cost X thousand pounds. More complex regression models would take into account more variables, such as square footage, location, distance to transport and then calculate the price accordingly. Machine Learning, and particularly neural networks, allow a machine to take into account huge numbers of variables and multiple levels of non-linearity to generate much more accurate predictions. The machine can process vast amounts of data rapidly, classify and then score information enabling a whole new class of automation.

"AI and Machine Learning are being used to build more effective financial crime prevention and detection tools. At ComplyAdvantage we are primarily known for our future forward approach to AML compliance software, applying AI to help identify and classify global money laundering risks for companies."

■ **Charlie Delingpole**, Founder & CEO of ComplyAdvantage

Creating effective AML data

High quality AML risk data is the lifeblood of any AML software solution. Insight into money laundering risks is inherently difficult to spot and dynamic in nature as it is contained within vast numbers of largely unstructured data sources spread across the world and is constantly changing day by day. AML risk data providers face a major challenge to effectively identify, classify and monitor these risks and the web of connections between entities and then provide this in an easily digestible format for organisations to use in their internal processes.

The traditional way of building a 'risk database' is to use large teams of analysts to review news and data sources to create structured profiles of individuals they find, researching each one individually. However, this has some significant inherent challenges. For example, if a provider has a database of 2-3 million profiles and claims to have 250 analysts updating 40,000 profiles a month, it would take many years to review each entity. Similarly, the analysts are limited by their individual language capabilities and the number of minutes allocated to the task. Consequently, the data can lack coverage and depth, resulting in, at best, a compliance professional having to conduct their own manual research (often 'Googling' the entity in question), or, at worst, missing risks entirely.

Applying AI to the creation of AML risk data can prevent compliance breaches by spotting previously unknown risks, updating entities faster, identifying remote linkages between entities and enhancing existing profiles with more information to help make better decisions more rapidly.

Machine Learning techniques like 'clustering' enable systems to make sense of vast quantities of data. Clustering works by creating clusters of entities, such as people or events, extracted from a series of images, text or audio. By grouping different groups of entities together based upon a range of characteristics, it is possible to distinguish those that are high risk or low risk, based upon entities that have been previously been flagged as high risk. The main aim of all clustering techniques is to form groups of similar entities based upon data points. This allows machine algorithms scanning sources of data to accurately spot risks indicated by the type of language used and the context of the passage to automatically classify them into a database, tagging all the relevant risk indicators and scoring the risk stage, age and type. The algorithms that search data sources for information learn and improve over time from repeated analysis of training data - data where the risks have already been confirmed by humans. They can do this until they can match and exceed the quality of human analysts.

Imagine new groups of entities are identified that are considered high risk. The risks can be identified quickly by assessing clustered linkages of high risk entities, linking them to specific risks such as illicit trade. By extracting the relevant features of the new clusters, you will be able to determine if the existing database of risk entities is connected to these new risk clusters. AI enables you to analyse each entity to find such associations within seconds, where it would take analysts many years to do the same. Similarly, if new entities are added, all the inter-connections with the existing dataset can be immediately analysed and refreshed.

Machine Learning enables systems to be more effective at extracting relationships between entities, which makes it possible to store data in a more efficient way such as using graph databases. These differ from 'flat file' databases (e.g. a typical spreadsheet) by storing information in a graph structure which uses nodes and edges to structure the data in such a way that links between the data can easily be found. Graph databases, by design, facilitate the simple and fast retrieval of complex relationships and links by using a greater array of tags which are applied to the data when it enters the database.

Screening for AML risk - onboarding, ongoing monitoring & payment screening

In today's globalised environment, matching a customer's identity with the identities in a risk database is hard work. Names can be written in different scripts, there can be regional variations in name spelling, word order can be different, cultural variations complicate matters further (e.g. inclusion of multi-generational names). Furthermore, there are legitimate errors to watch out for, as well as deliberate attempts to deceive the system and firms are often also constrained by siloed and/or poor quality internal data.

When a human compares a potential match, they rely on a great deal of contextual information to make a judgement call. Where it is difficult to determine a match, there will always be a grey area and the investment to improve confidence in a decision will be driven by an organisation's risk-based approach. The traditional screening process is often highly inefficient, generating lots of unnecessary alerts, disrupting the customer experience (e.g. holding up payments unnecessarily) and requiring a huge overhead of manpower to process.

One reason traditional systems struggle is that they typically rely on static rules-based name matching criteria to identify potential name matches along with the use of secondary identifiers (e.g. date of birth, location) to reduce false positives. A human is then relied on to review the alerts to check the 'context' and make a decision.

Further complicating matters, organisations often periodically ingest a 'flat file' of a provider's entire AML database and use a third-party solution to query it, which means they are never querying an up to date version of the data. In addition, the search algorithms are sub-optimal as they were not built around that specific data set. Even worse, a lack of faith in any one data provider often means that multiple data sets are used, creating duplicate profiles further increasing the 'noise'. As a result, the types of risk signals screened for and the 'fuzziness' of the search is often limited to 'goal seek' the number of alerts the team can manage from an operational perspective.

AI enables a system to use linguistic search technology to shift from a 'name match' towards using a contextual driven approach of 'identity matching'. A complex data and search challenge that requires the processing of vast amounts of data in a short time frame (e.g. with the shift towards real-time straight-through payments) is perfectly suited to AI and Machine Learning techniques that can significantly reduce false positives and false negatives.

Machine Learning can be used to take into account a huge number of attributes and take a holistic view to score the probability of a true identity match, rather than just flagging when any one linguistic rule is broken. Given the nature of a graph database, Machine Learning-powered search algorithms can be used to query the linkages between entities far more effectively and efficiently than a rules-based system analysing a flat file.

AI-enabled systems can also learn from users' decisions and proactively suggest changes in the rules to improve the relevance of alerts. For example, improving the targeting of which parts of an AML risk data set are relevant to an organisation (by source type, crime type, risk age, geography, etc.) to empower a compliance team to continually optimise their use of the system.

AI is also able to present a suggestion of the most likely decision the user will make for lower risk scenarios – whilst highlighting the most relevant information and the rationale behind the recommendation to further empower the user to prioritize their time on more serious, higher probability and more complex issues – which is where humans really add value.

Identifying suspicious behaviour - transaction monitoring

Transaction monitoring involves analysing internal data (e.g. payments) to identify potentially suspicious behaviour, at which point a firm is required to investigate and either create an internal record or file a report with the regulator. As with AML screening, this is a data and analytics challenge that today often generates a high volume of mostly unnecessary alerts for analysts to manually review in an effort to spot the real risks.

Historically, this has been done by retrospectively processing batches of payment messages against a small number of behavioural rules, adjusted by customer segment. Efficiency is limited by a lack of flexibility in the rules (which may make sense at a segment level, but not for some individual clients within that), and too narrow a data focus (lacking the context of the broader customer relationship).

AI and Machine Learning can be utilised to improve these processes in many of the ways discussed in relation to AML screening, such as improved analysis of linkages and proactive suggestions to improve rule effectiveness.

Whereas static rules often rely on perfect internal data (which rarely exists) to work efficiently, Machine Learning enables the use of more context to work around the available data to create a suitable risk score. The scoring can also be much more refined as the system can analyse a much higher number of contextual data attributes, for example taking into account profile information (e.g. from a customer relationship management (CRM) system), other behavioural information beyond payments (e.g. website login activity) and other unstructured data sources.

Unsupervised learning techniques can also be employed to analyse historical data sets and proactively suggest anomalies that wouldn't normally be identified using a solely rules-based system, providing further value.

Given many of these techniques are increasingly being adopted in other areas of financial crime such as fraud identification and market abuse, it seems likely that we will see the same applied more extensively in transaction monitoring, empowering and complementing the regulatory requirements around rules-based scenario monitoring, audit trails and reporting.

Realising the promise of Artificial Intelligence

As we have seen, organisations face a major challenge to prevent money laundering whilst improving efficiencies and reducing costs. PwC estimates that between 90-95% of all AML risk alerts are false positives. [2] Remediating the majority of these is a time-consuming and robotic manual process that requires a high level of concentration to avoid errors. Moreover, pressure from senior staff to deal with this revenue draining task can result in individuals 'sweeping under the carpet' risk alerts that would take too long to investigate.

At the core, many of the underlying challenges are driven by the need to analyse large quantities of data extremely rapidly. AI and Machine Learning are techniques that can improve the ability of a software solution to execute these tasks at scale and speed whilst more closely replicating the thought process of humans and enabling an automation of lower level, manual tasks. By improving the ability to spot risks and reduce overheads, AI and Machine Learning can enable businesses to concentrate their human capital on higher priority risks - feeding into the Financial Action Task Force's (FATF) emphasis on utilising a risk-based approach to compliance.

AML solutions powered by AI and Machine Learning can improve risk identification and dramatically reduce the number of false positives – at ComplyAdvantage we often see reductions from 60-80% and above. As previously mentioned, improving the quality of the AML data is a major contributor, supported by more sophisticated screening and monitoring algorithms. Similarly, when alerts occur, richer profiles coupled with an intuitive user interface can help analysts remediate alerts more quickly.

In our experience, as well as reducing costs, using AI and Machine Learning can enable firms to screen/monitor for more risk signals. For example, when screening payments, firms often screen beneficiaries against the bare minimum of entities on a few core sanction lists at a very low level of fuzziness to minimise hits whilst ticking the compliance box. Effectiveness is therefore severely limited (your customer is unlikely to send money to someone whose name is on a sanction list, but when they send it to that person's relative, will your system pick it up?) With legacy 'noise' removed, organisations can widen the net and screen for more risk signals, e.g. reputational risks such as sending money to terrorists who are in the press but will take months to make it to the official lists.

Final thoughts

Preventing the formal financial system from being used for money laundering is a considerable challenge. As criminals become more sophisticated and use an increasingly diverse range of money laundering methodologies, it is hard for the private sector to stop them using conventional AML solutions. Given these challenges, it is essential that companies give their compliance teams the appropriate tools for the job. In a world of ever increasing data, AI and Machine Learning-powered systems can be used to swing the balance back in favour of the good guys.

Notes

1. https://www.merriam-webster.com/dictionary/artificial%20intelligence
2. https://www.pwc.com/us/en/anti-money-laundering/publications/assets/aml-monitoring-system-risks.pdf

■ **Charles Delingpole** is CEO and founder of ComplyAdvantage (+44 (0)20 7834 0252, contact@complyadvantage.com).