MUST READ: Colonial Pipeline paid close to $5 million in ransomware blackmail payment

Big Data versus money laundering: Machine learning, applications and regulation in finance

Could financial fraud such as the Laundromat be avoided by applying machine learning to scan through data? And if yes, why is that not happening?

By George Anadiotis for Big on Data | March 23, 2017 -- 14:18 GMT (07:18 PDT) | Topic: Innovation

Predicting and acting upon financial fraud is one of the prime areas of application of advanced big data techniques like machine learning (ML). Earlier this week, a case of money laundering known as the Laundromat was uncovered (https://www.theguardian.com/world/2017/mar/20/british-banks-handled-vast-sums-of-laundered-russian-money) by the Organized Crime and Corruption Reporting Project (OCCRP) involving a number of global banks active in the UK.

Could ML help prevent such incidents? What progress is there on this front, how does it fit in the bigger picture, what are the roadblocks, and what may be the repercussions of adoption?

"It isn't just individual transactions. It's the repeated pattern"

SPECIAL FEATURE

(/topic/internet-of-things-the-security-challenge/)

IoT: The Security Challenge (/topic/internet-of-things-the-security-challenge/)

The Internet of Things is creating serious new security risks. We examine the possibilities and the dangers.

Read More (/topic/internet-of-things-the-security-challenge/)

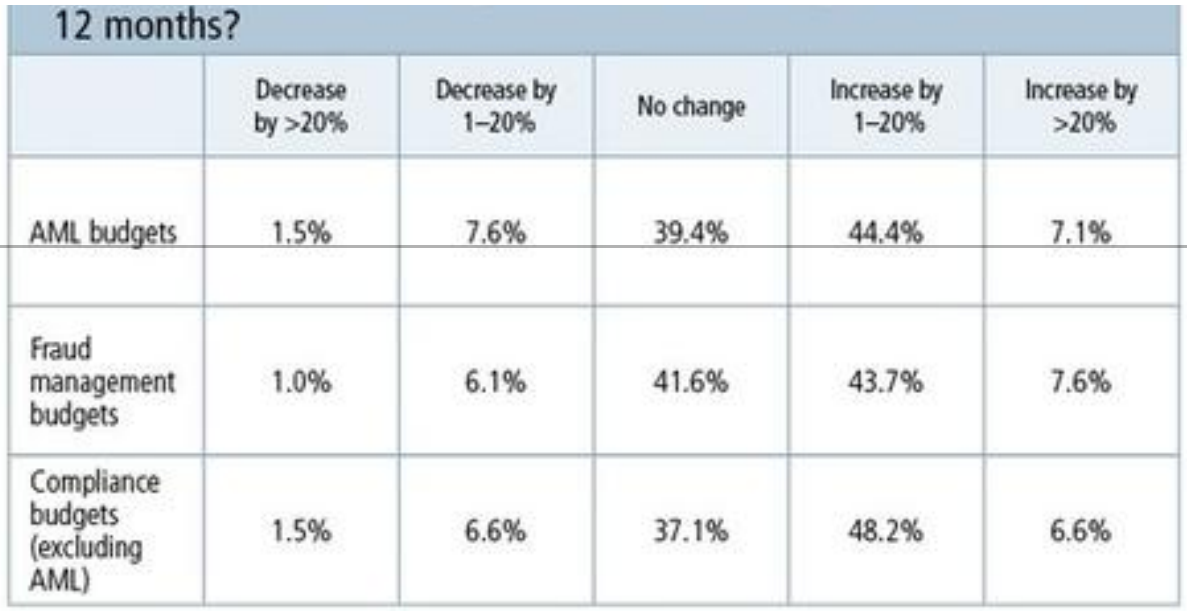
There are many different types of fraud related to the financial industry. The Laundromat is a case of money laundering (MLA), which is estimated to generate about US$300 billion in illicit proceeds annually in the US alone.

While each type of financial fraud has its own characteristics and implications, MLA is considered important enough for the US to have its Department of the Treasury produce a National Money Laundering Risk Assessment (https://www.treasury.gov/resource-center/terrorist-illicit-finance/Documents/National%20Money%20Laundering%20Risk%20Assessment%20-%2006-12-2015.pdf) (NMLRA) report in 2015.

The reason MLA carries this weight is clear even without reading the 100-page long document in its entirety. MLA has more than financial impact, as it is associated with activities ranging from trafficking people and drugs to terrorism and corruption. It's no wonder then that governments around the world are trying to crack down on MLA by means of regulation on financial institutions.

Financial institutions have to comply with a set of rules imposed by regulators, and are audited to verify their compliance. If found in negligence of their duties, they are faced with legal consequences. For example, HSBC-US entered into a deferred prosecution agreement (DPA) in the US in 2012, for failing to adequately monitor more than US$670 billion in wire transfers and $9.4 billion in purchases of U.S. bank notes from HSBC Mexico.

1 Given the current macroeconomic environment, how are financial crime and compliance budgets expected to change over the next



12 months?

	Decrease by >20%	Decrease by 1–20%	No change	Increase by 1–20%	Increase by >20%
AML budgets	1.5%	7.6%	39.4%	44.4%	7.1%
Fraud management budgets	1.0%	6.1%	41.6%	43.7%	7.6%
Compliance budgets (excluding AML)	1.5%	6.6%	37.1%	48.2%	6.6%

Anti-Money Laundering is serious business, and this is also reflected in how organizations are forecasting resources they are going to invest in it. Image: BAE Systems and Operational Risk

It's no wonder then that financial institutions appear in their turn to be taking anti-MLA compliance seriously: 51.5 percent of respondents in a recent survey (https://www.risk.net/risk-management/operational-risk/2467001/financial-crime-survey-2016-compliance-and-online-fraud) drawn from banks and insurers who work in risk, fraud, compliance and finance said that anti-MLA budgets would increase. But is this money well-spent? Judging from the HSBC example, maybe not so much.

According to the OCCRP, HSBC is the main culprit in the Laundromat case, having processed more than US$500m in cash through its British and foreign branches. Banks like HSBC claim that despite having sophisticated units dedicated to rooting out financial crime, the volume of payments -- billions a year -- makes such work difficult.

Others, like L Burke Files, an international financial investigator, call compliance checks at many western banks "desultory, and often little more than box ticking." Files however also notes: "Most of the transactions I'm seeing here would have required substantial enhanced due diligence. It isn't just individual transactions. It's the repeated pattern."

Rules are a blunt instrument, machine learning is a black box

Repeated patterns and transaction volumes in the billions? This sounds like a job for ML. Sunil Mathew is the head of the Financial Crime and Compliance unit in Oracle Financial Services (https://www.oracle.com/industries/financial-services/banking/index.html) (OFS), and his job is to work with 9/10 major banks worldwide to help them comply with anti-MLA regulations. Part of that is looking into the applicability of ML in this domain.

OFS works with their clientèle to look at the banking products they have, the markets in which they operate and the regulations that apply in those markets to understand the risks they try to address. Then they map these risks to controls that need to be in place, and provide detection scenarios that implement these controls.

Mathew notes that in the last 15 years a set of commonly accepted scenarios has emerged for regulators around the world. One of those scenarios is monitoring rapid movement of funds as an indication that may point to MLA and generate alerts.

But even though the broad scenario may be the same, its parameters will vary: the volume of funds to monitor, the rate and time window of movement and the risk profiles of parties in the transactions to be monitored are some of these parameters.

Oracle ships such scenarios as part of its products that users can customize according to their needs. This rule-based approach works, but as Mathew puts it, "rules are blunt instruments. They may trigger to catch bad guys, but they will trigger for many good guys too." This is a problem as it means that the people whose job is to check on those alerts will have a bigger workload, and it's the reason that Oracle is incorporating ML in its products.

ML algorithms are a good match for this scenario, as they can use training data to be developed and then customer-specific data to be fine-tuned, resulting in higher accuracy and increased performance.

Although Mathew was not able to share results, ML approaches used today in domains like speech recognition are known to be able to achieve accuracy in the area of 95 percent (https://www.zdnet.com/article/ibm-vs-microsoft-human-parity-speech-recognition-record-changes-hands-again/). There is one problem though: ML is, as Mathew puts it, a black box.

When used to determine how banks will market their products or what offers they will make to their clients, this is not so much of a problem -- regulators do not care about how these processes work. But when it comes to compliance, showing results is not enough: banks need to be able to explain how they arrived at those results.

This is one of the key challenges with ML: "The more sophisticated algorithms are essentially a black box, and you can't open the box to look what's inside. This has been a major roadblock

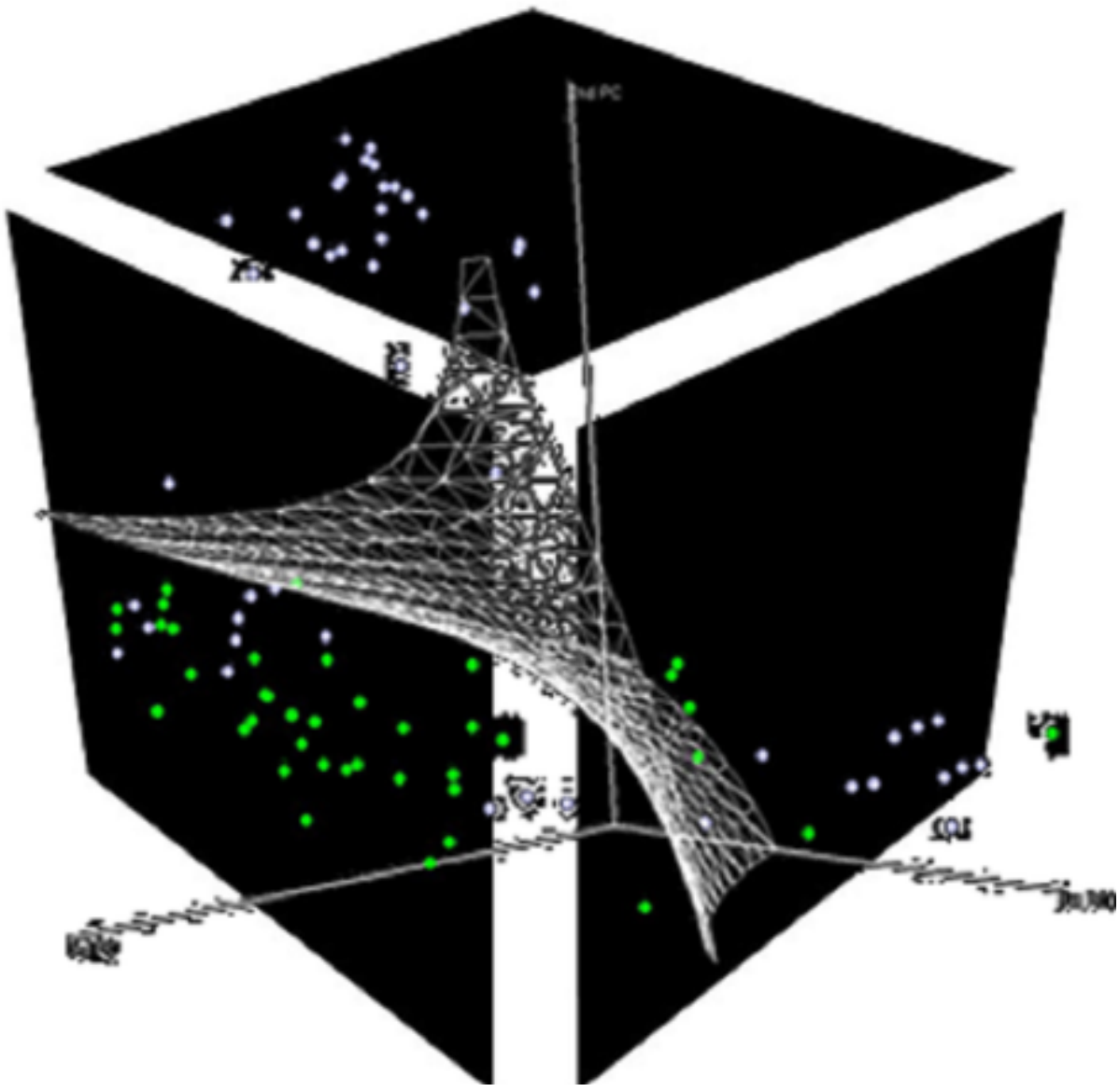
for adoption," says Mathew. But the stakes for Oracle and banks are too high to give up on ML, so they are trying to apply different approaches to tackle the issue.

Building trust in the black box

The first approach is pragmatic and directly applicable: if regulators are not comfortable with accepting ML as the core of the anti-MLA engine, keep rules as the core and apply ML to evaluate generated alerts.

By training ML on the course of actions taken on alerts, they can identify patterns that help classify them as more or less likely to signify MLA, helping prioritize them. This is in line with the tendency to progressively inject advanced functionality in organizations to support everyday operations (https://www.zdnet.com/article/artificial-intelligence-in-the-real-world-what-can-it-actually-do/).

The second approach is also pragmatic, although more forward-looking: trying to work around the concerns of regulators. Oracle is working with some of its clients to convince



regulators that ML can be used as a tool for anti-MLA.

What can be done to open up the black box of machine learning? Image: aitrends

The reasoning is that even though you cannot see inside ML models, by building enough controls around them and independently testing and auditing them it should be possible to verify that they work as they are supposed to. Having the data that algorithms work on is an integral part of it as well.

The third approach is working on removing the barrier altogether: "In Oracle we are lucky enough to be working with our research labs, and one of the areas we are focusing on is making ML more interpretable" says Mathew.

As anti-MLA is in many ways about connecting dots, this naturally lends itself to a graph processing paradigm. Graph processing has been used in cases such as exploring connections in the Panama Papers data (http://www.odbms.org/blog/2016/10/how-the-11-5-million-panama-papers-were-analysed-interview-with-mar-cabra/), and a mixed approach utilizing both ML and graphs may produce results of broader interest.

ML is efficient, but opaque: "It works, and it works well, but we do not exactly understand why or how (https://www.zdnet.com/article/artificial-intelligence-in-the-real-world-what-can-it-actually-do/)." Although that has been said on deep learning, it applies more broadly for ML as well, and coming from experts in the field it is not something to be dismissed lightly.

This may raise some philosophical questions, mostly having to do with the increasing feeling of being sidelined and not being able to keep up with technology, but there are also some very practical implications.

As Mathew notes, whatever anti-MLA approach taken, getting results is not enough. It must also comply with a number of guidelines, ensuring for example there are no discriminations against certain groups of the population.

The issue of algorithmic transparency is becoming increasingly understood and widely discussed, and there are many examples in which opaque algorithms embody all sorts of bias

(https://gotocon.com/berlin-2016/presentations/show_talk.jsp?oid=7974). If regulators decide to adopt ML in the financial industry, it will be interesting to see under what conditions it will be done and what will be the repercussions.

The human factor: efficiency versus transparency and distribution

But the question of whether opaqueness is a price we are willing to pay for efficiency is not the only one here: does regulation work? And who will benefit the most from adopting advanced techniques in the finance industry?

Mathew points out that although Oracle may sometimes brief regulators, it's the banks that work with them. Apart from the obvious question on the relationship between them (https://dealbook.nytimes.com/2013/05/23/banks-lobbyists-help-in-drafting-financial-bills/), do regulators have the resources and knowledge to keep up in this arms race of sorts?

The UK for example has been hailed by UBS as having "progressive regulation and established support for new innovations (http://www.businessinsider.de/ubs-launches-robo-advice-product-smartwealth-2016-10)" but how does that translate? Mathew notes that in areas like capital management, some regulators have advanced knowledge of statistical techniques and predictive models, but compliance is different.

It looks like even though the future is here, it's not evenly distributed (https://medium.com/not-evenly-distributed/the-future-has-arrived-fed56cec3266#.bs8kc89fx): the expectation is that innovation will eventually even things out (https://www.youtube.com/watch?v=1vr6Q77IUHE) and create new jobs, but whether or when that will happen, or what happens in the meanwhile, are open questions.

In a world of increasing inequality (http://www.oecd.org/social/inequality.htm), is technological innovation making society more unequal (https://unu.edu/publications/articles/is-technological-innovation-making-society-more-unequal.html)? This is an ongoing debate, and the financial industry is a striking example of applying innovation to amass unevenly distributed wealth.

"Banks have a desire to move to modern techniques, because it will save them money. Some banks may have up to 6,000 people working on compliance," says Mathew. While it is obvious that automating a tedious, error prone and time consuming task like anti-MLA will bring great benefits, people for which anti-MLA is their job may have a different view on this.

These are all questions that go far beyond anti-MLA. The finance industry is not only a good example of how big data innovation can be applied, but also of the implications that come as part and parcel of this. As Mathew says, "this is an exciting area, and things are just starting to happen."

INTERNET OF THINGS

(https://www.zdnet.com/article/who-really-owns-your-internet-of-things-data/)

Who really owns your Internet of Things data? (https://www.zdnet.com/article/who-really-owns-your-internet-of-things-data/)

In a world where more and more objects are coming online and vendors are getting involved in the supply chain, how can you keep track of what's yours and what's not?

Read More (https://www.zdnet.com/article/who-really-owns-your-internet-of-things-data/)

RELATED TOPICS: BIG DATA ANALYTICS | CXO | DIGITAL TRANSFORMATION | TECH INDUSTRY | SMART CITIES | CLOUD

By George Anadiotis for Big on Data | March 23, 2017 -- 14:18 GMT (07:18 PDT) | Topic: Innovation

SHOW COMMENTS

HOURS AFTER ATTACK

Colonial Pipeline paid nearly $5 million to ransomware gang

JUST IN

The best home office cooling in 2021: Solutions to prevent overheating

1 hour ago

Two in three parents would quit their jobs without work from home flexibility

2 hours ago

How to get rid of the Gmail Meet tab on iPhone and Android

2 hours ago

In smartphone exit, LG's loss will be Samsung's gain

2 hours ago

Colonial Pipeline paid close to $5 million in ransomware blackmail payment

3 hours ago

How hyperscalers will allay Musk's Bitcoin energy consumption worries

3 hours ago

Microsoft's new security feature locks hackers out with GPS

3 hours ago

Alibaba's cloud growth hits slowdown after loss of top customer

3 hours ago

TODAY ON ZDNET

SPECIAL FEATURE

Analytics: Turning data science into business strategy

Quantum computing: Intel's cryogenic chip shows it can control qubits even in a deep freeze

3 hours ago by Daphne Leprince-Ringuet in Quantum Computing

The best free music services in 2021: Top free streaming apps and radio picks

3 hours ago by Jason Perlow in Mobility

Walmart announces plans to purchase virtual fitting room company Zeekit

4 hours ago by Jonathan Greig in E-Commerce

Microsoft just achieved something few ever thought possible

4 hours ago by Chris Matyszczyk in Microsoft

The Chromebook turns 10: Cheap, safe, powerful -- and still gaining on Windows

4 hours ago by Steven J. Vaughan-Nichols in PCs

Why fluid business conditions require liquid networks

15 hours ago by Enzo Cocotti in Networking / Paid Content

VIDEO

The psychology of cybersecurity: How empathy is key to keeping people safe online

Rushing into AI? Read this book first

4 hours ago by Tonya Hall in Artificial Intelligence

AlmaLinux checklist: 9 things to do after installation

from TechRepublic Premium

PayPal taps Google Cloud for more infrastructure, analytics services

4 hours ago by Stephanie Condon in Cloud

Pure Storage updates Portworx Enterprise with new integrations

4 hours ago by Stephanie Condon in Storage

AI in sixty seconds

5 hours ago by Tiernan Ray in Artificial Intelligence

The best cybersecurity certifications in 2021: Deepen your knowledge

5 hours ago by Charlie Osborne in Security

GALLERY

CES 2021: Best of Innovation award winners

LOAD MORE

Collection

Coronavirus: Business and technology in a pandemic

Best home office cooling solution 2021: Prevent overheating

Best free music service 2021: Top streaming apps and radio

Best cybersecurity certification 2021: Deepen your knowledge

Best music service 2021: Premium music streaming apps

Best laptop for college 2021: Notebooks for students

Best Android app for power users 2021: Track your data usage

Collection

Working from home: The future of business is remote

Colonial Pipeline paid close to $5 million in ransomware blackmail payment

Microsoft's new security feature locks hackers out with GPS

Best cybersecurity certification 2021: Deepen your knowledge

Ransomware: How the NHS learned the lessons of WannaCry to protect hospitals from attack

Microsoft warns: Watch out for this new malware that steals passwords, webcam and browser data

Fake Android, iOS apps promise lucrative investments while stealing your money