Guidehouse

Inside the Black Box: Artificial Intelligence Applied

Inside the Black Box: Artificial Intelligence Applied

This article was originally published in Money Laundering Bulletin

More than £100bn a year is estimated to be laundered through the UK financial system, and UK financial institutions spend an estimated £5bn a year on financial crime prevention. Most of this is spent on transaction-monitoring systems, sanctions filters, Know-Your-Customer (KYC) and Customer Due Diligence (CDD) systems, and the recruitment and retention of compliance staff to detect suspicious activity. Firms are starting to leverage artificial intelligence and machine learning (AI/ML) to focus on transactions that present real risk by producing more effective alerts and reducing the volume needing human intervention.

In April 2018, HSBC deployed AI/ML to detect money laundering, terrorist financing, and fraud in customer accounts. Jennifer Calvery, HSBC's global head of financial crime threat mitigation, said the bank believed in "harnessing technology and data ... to get to a place in the future where we understand and can see criminal behaviour in as nearly real time as possible". HSBC has partnered with RegTech startup Ayasdi to implement an AI/ML solution that will more effectively identify suspicious activity and reduce the number of false-positive alerts generated by traditional transaction-monitoring and sanctions-filtering systems.

Guidehouse has also partnered with Ayasdi to deliver transaction-monitoring solution services to financial institutions headquartered in Europe and the U.S. In 2018, Guidehouse was engaged by a large European financial institution to deploy machine intelligence in a four-year Anti-Money Laundering (AML) look-back of correspondent banking activity, using intelligent segmentation and intelligent typologies to target the highest-risk areas.

How does intelligent segmentation work?

The traditional segmentation of the customers uses available KYC data, e.g., whether a customer of the bank is a corporation or a financial institution, along with the products/services the customers have signed up for, and other information to predict their future behaviour. The customers sharing similar information are typically grouped together (in "segments") and the shared information is used to predict the type, frequency, and value of transactions, as well as the counterparties the customers are likely to transact with. Transactions occurring outside the expected parameters generate the transaction-monitoring alerts.

Guidehouse and Ayasdi's intelligent segmentation relied instead on the past behaviour of the customers to predict their future behaviour. This process looked beyond information declared on

1. Transparency International [2018], UK's "£100 Billion" Dirty Money Problem Must Be Confronted Following International Review, available at https://www.transparency.org.uk/press-releases/uks-dirty-money-problem-confronted/.
2. Chris Stokel-Walker [2018], Why the UK is losing its costly battle against money laundering, Wired, available at https://www.wired.co.uk/article/money-laundering-in-the-uk-russian-banks.
3. Martin Arnold [2017], HSBC brings in AI to help spot money laundering, The Financial Times, available at https://www.ft.com/content/b9d7daa6-3983-11e8-8b98-2f31af407cc8.

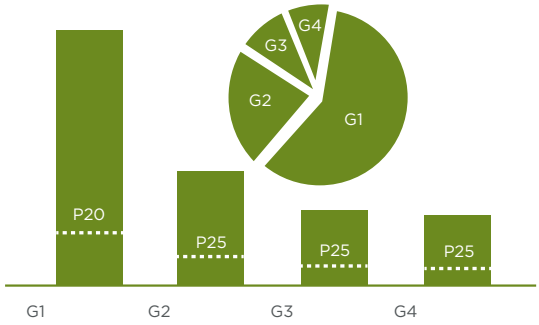
KYC forms. The ML algorithms were fed data points available on past transactions (e.g., transaction value/amount) as well as values derived from these data points, known as features (e.g., the average number of transactions in a month involving four-plus countries[4]). The algorithms discovered new segments of customers. Many of the new segments were more granular than the traditional segments and others transcended multiple traditional segments.

How did this help the financial institution?

The better granularity allowed the transaction-monitoring thresholds to be tuned more accurately and the parameters for detection scenarios to be set more tightly around expected behaviour. Furthermore, newly discovered segments allowed some transactions to generate new alerts that were previously overlooked. Intelligent segmentation reduced the alert population by 45% using segmented tuning, and generating up to 15% more productive alerts. Thus, the financial institution was able to increase both the effectiveness and efficiency of its transaction-monitoring process because investigators examined fewer alerts that were statistically more likely to represent a real money laundering risk.

Below is a Traditional Segmentation vs. Intelligent Segmentation Diagram:

A **typical segmentation process** produces unevenly distributed groups, or may rely on a single data point for segmentation.

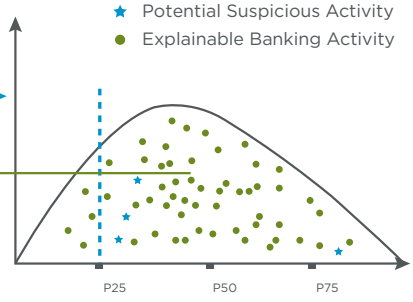


To alert on the potential suspicious activity of this segment, a lot of explainable behavior also needs to be reviewed.

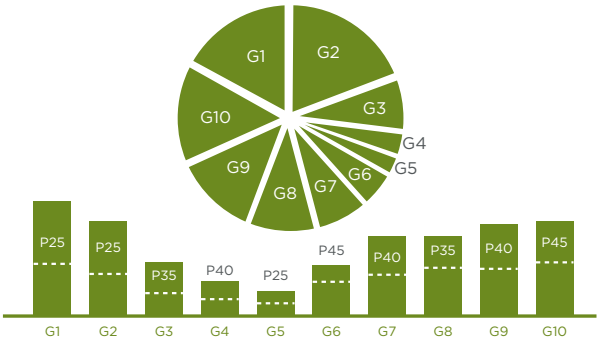
Visualizing the Process for One Segment Group

Rule Segments

Segment	Threshold
Retail	T1
Corporate	T2
Commercial	T3
Small Business	T4



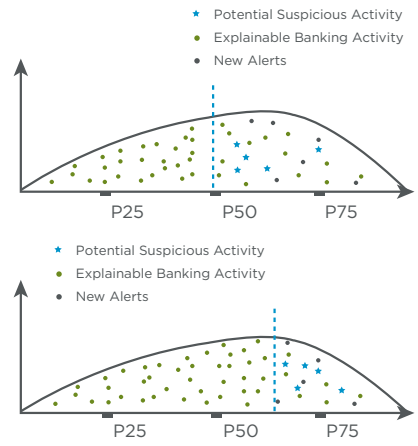
Intelligent segmentation identifies segments using a collection of related data points beyond traditional identifiers to identify groups that should be monitored together and those that should likely be split.



Visualizing the Process for Two Segment Groups

New Rule Segments

Scenario	Threshold
S1	T1 - New
S2	T2 - New
S3	T3 - New
S4	T4 - New



4. In this scenario, we have taken four plus countries to include correspondent / intermediaries bank locations.

How could you incorporate AI/ML into your AML processes?

AI-based systems are more robust than rule-based software and can adapt far more quickly to changing money laundering patterns. Vast amounts of structured and unstructured data can be ingested to identify complex criminal activity across different products, lines of business, and customers. RegTech companies have developed segmentation, clustering, profiling, modelling techniques, and statistical analysis to detect financial crime and reduce false positives by up to 30%.[5]

Below are several examples of how firms can harness AI/ML-enabled solutions to improve both the effectiveness and efficiency of their anti-financial crime processes. In each example, we see organisations taking a phased approach to AI/ML implementation, i.e., starting with tools that learn how the firm handles alerts, moving on to prioritisation of alerts/issues, quality assurance, and finally to taking over tasks previously performed by humans.

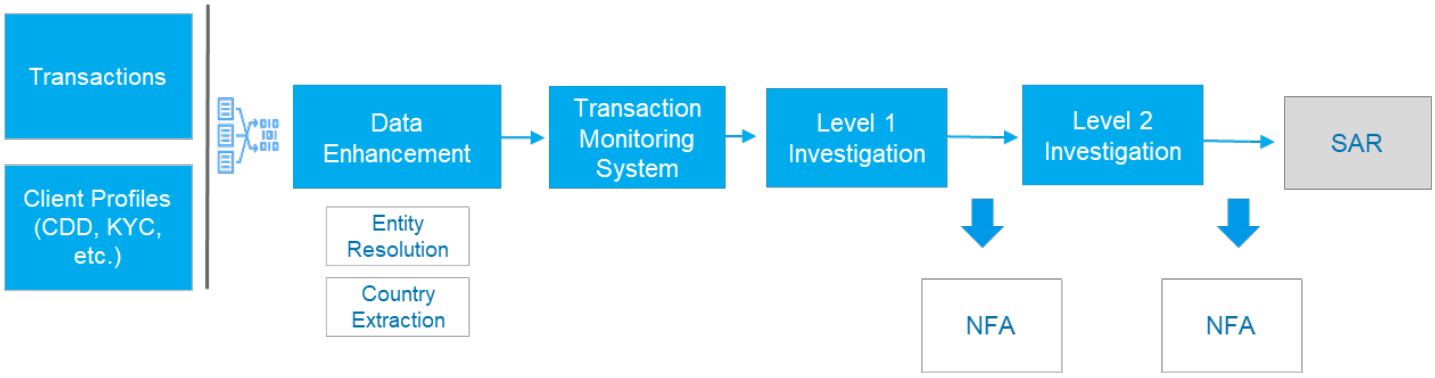
Transaction monitoring

Traditional transaction-monitoring systems depend on rule-based scenarios that require continuous tuning. Costly and time-consuming, the tuning puts a financial institution at risk of missing newly devised money laundering schemes. AI/ML can reduce the number of false positives by taking more data into account to identify complex criminal activity across different products, lines of business, and customers.

Case Study No. 1

The problem:
As we've seen above, the problem with even the most sophisticated traditional AML transaction-monitoring systems is that customers are segmented based on known attributes that don't necessarily dictate behaviour. The transaction-monitoring rules are tuned based on the behaviour of the segment. Therefore, the net is cast too wide to detect anomalous or suspicious behaviour.

Below is a diagram of a traditional transaction-monitoring system life cycle. Please note that in the following diagrams SAR refers to Suspicious Activity Report and NFA means No Further Action:



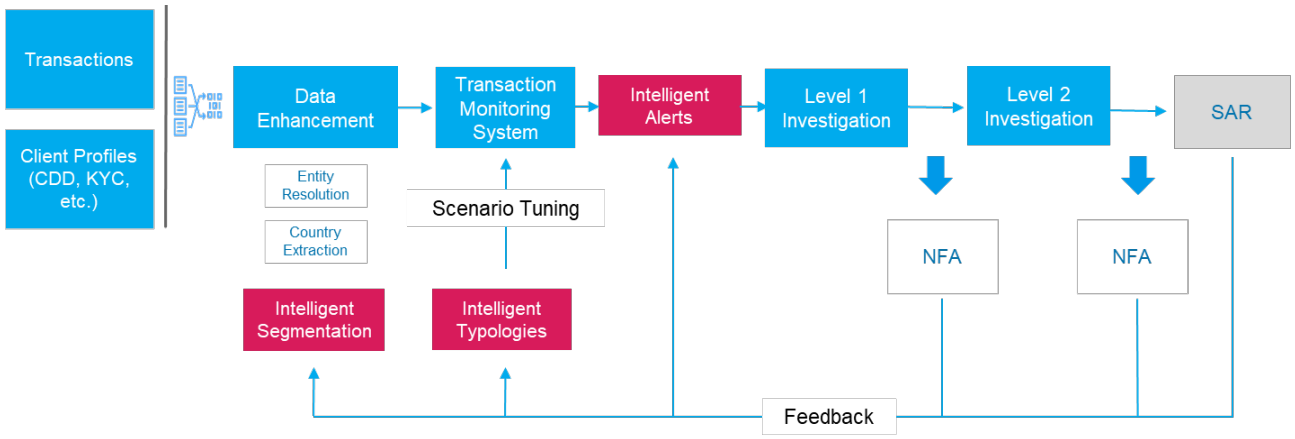
5. McKinsey & Company [2017], The new frontier in anti-money laundering, available at https://www.mckinsey.com/business-functions/risk/our-insights/the-new-frontier-in-anti-money-laundering.

Solution 1: Augmented AML TM Process

An AI/ML-enabled solution can enhance the traditional transaction-monitoring system by augmenting the:

- Data-segmentation process for more effective and efficient thresholds;
- Alert-prioritisation process for improved alert investigation and resolution; and
- Risk typology-identification process for the development of enhanced scenario logic.

The below diagram illustrates an AI/ML-enabled transaction-monitoring solution:

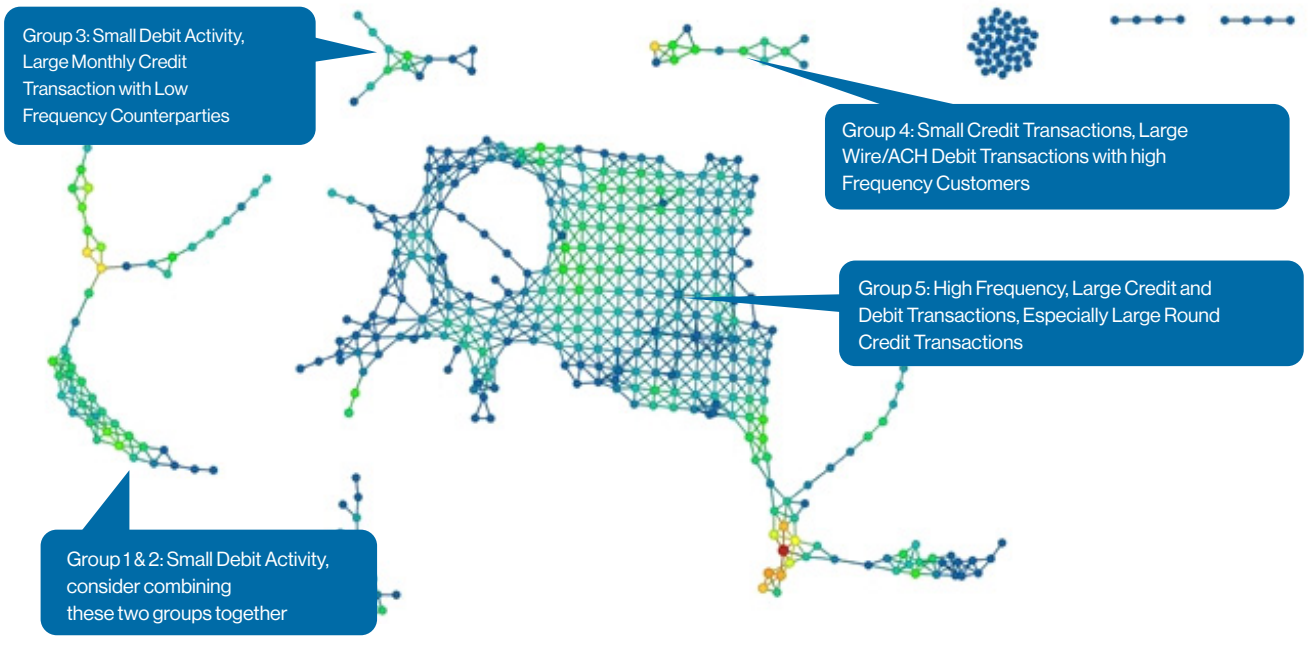


Solution 2: The Topological Data Analysis (TDA) Model

TDA creates families of groups using a topological model. Topology is the study of shape. The shape tells you the fundamental "structure" of the data and expresses instantly what the data is trying to "say".

The value of TDA is that there is no need to define the parameters of analysis. This means that identified groups can be based on any number of varied attributes – behaviour, customer type, investigation outcome, etc. Below is an illustrative diagram:

Intelligent Segmentation



Sanctions filtering

Implementing an AI/ML-based solution into existing sanctions-filtering systems can significantly reduce the volume of alerts requiring human intervention, while maintaining a risk-averse position. Under a traditional sanctions-screening tool, when a threshold is reached (e.g., an 80% likelihood that the name or address in the payment matches a watchlist subject) an alert is generated that requires a human to designate whether the alert matches the watchlist entry. Up to 98% of these alerts are obvious false matches[6] generated by the fuzzy matching capabilities of the screening tool. These false matches can be disposed of using only the information available on the review screen.

How does traditional sanctions screening work?

The fuzzy matching of the screening tools, in a nutshell, compare two strings of text characters (e.g., names) and match them if they are approximately the same according to a pre-set threshold of similarity. Some of the most ubiquitous methods to calculate this similarity are based on the Levenshtein Distance, its variants or extensions. This is illustrated in the box on the right.

LEVENSHTEIN DISTANCE EXAMPLE

1st string: JILL WHITE BASS
2nd String: BILL WIHTE BAS

Levenshtein distance = 4, which consists of 3 units of distance due to 3 replacements and 1 unit of distance due to 1 deletion

Length of the longer string = 13, not counting spaces - the longer of the two strings is the 1st string in the example

Match score = $100\% - \left(\frac{4}{13}\right) = 100\% - 30.8\% = 69.2\%$

If an institution pre-sets the similarity score at 65%, then an alert would be generated in this example.

When adjudicating the alerts, the human reviewers use information available on the review screen, from the payment transaction data being screened or the profile of the watchlist subject, to determine whether the alerts are indeed true or false matches. Several things factor into this decision-making other than the string similarity. For example, a gender mismatch shown in the JILL vs. BILL example above may persuade the alert reviewer to flag the match as a false positive. In other cases, two persons in entirely unrelated geographies could be the basis for a false positive adjudication.

How can an AI/ML-based system help?

An AI/ML-based screening solution makes use of contextual data points such as those described above, in addition to a singular string similarity score, in determining a match. The information derived from immediately available data points is commonly referred to as a "feature". In fact, several different string similarity algorithms apart from Levenshtein Distance (e.g., Jaro-Winkler) can be entered as features and considered by the AI/ML solution concurrently.

6. Patrick Angeles [2018], AML: Past, Present and Future – Part II, available at http://vision.cloudera.com/aml-past-present-and-future-part2/.

To mimic the human reviewers, the AI/ML-based solution is trained using real human dispositions of historical alerts to make sense of the various data points and features fed into it, including the similarity scores. The ML algorithms attempt many combinations, weighting, and ordering of these data points and features to create a pathway that most closely resembles the past human decision-making.

This trained AI/ML-based solution, when applied toward new alerts, can assist human reviewers in their dispositioning of the alerts or reduce human interventions at the stage of determining the most basic but also most numerous false positives.

Guidehouse was engaged by a North America-based global bank to implement a machine learning "proof-of-concept" for sanctions screening of transactions with the goal of making the AI/ML solution a pair of eyes in the bank's traditional "four-eye" alert review process. The goal here is to allow one of the two human employees to perform more-relevant tasks, such as actually investigating potential sanctions breaches.

With new data, further training, and tuning, the AI/ML solution has the potential to one day perform the entire "four-eyes" process with minimal human supervision. Indeed, it is not difficult to envisage that by integrating this AI/ML solution into the traditional filtering/ systems, alert generation can be reduced at the very beginning of the process.

The AI/ML solutions employ self-learning so that the same alerts that proved to be false positives are not generated again and again, unless something changes. The tool can also combine ML and natural language processing to screen accurately against any sanctions watchlist. This can also provide financial institutions with a clear understanding of the KYC profile of their customer's suppliers and employees, so the risk of sending inappropriate payments to blacklisted countries, companies, and individuals is significantly reduced.

KYC compliance

AI/ML can improve the KYC compliance process by reducing costly and time-consuming maintenance of the financial institution's KYC data. Traditionally, large financial institutions with branches around the globe need to utilise data from multiple sources and process documentation in disparate languages and formats to comply with KYC requirements. An automated KYC solution powered by AI and robotic process automation can ensure that customer documents are efficiently scanned, tagged by the AI solution, and uploaded to a central KYC repository.

The solution can also support the KYC system by providing constant monitoring of a customer's activity. The solution will automatically update the KYC repository when it identifies a change in the customer's profile via the institution's transaction-monitoring system or public domain research. This will reduce the burden on firms to manually monitor their customers' transactions and continuously search the web for adverse media. Instead, the resources can be redirected toward more important tasks, such as performing quality assurance.

Conclusion

The rapid evolution of new technology presents real opportunities for firms to gain an advantage in the global tech arms race with financial criminals. The UK Financial Conduct Authority's approach — encouraging innovation while ensuring firms continue to meet their regulatory obligations — should in the medium term position the UK in the forefront of international regulatory efforts to counter financial crime. At the same time, supervisors will want to understand the governance around the system, the risks it poses, and how they are to be mitigated. Firms should expect their supervisors to be particularly concerned with how the compliance function aims to avoid the "black-box-in-the-corner" phenomenon, where, over time, a technology becomes less and less understandable to its users.

Contacts

James Siswick
Managing Director
Global Investigations & Compliance,
London
M +44 (0) 20 7661 0570
E jamessiswick@guidehouse.com

Dave Bradshaw
Director
Global Investigations & Compliance,
London
M +44 (0) 20 7661 0624
E davebradshaw@guidehouse.com

Johnny Zhang
Associate Director
Global Investigations & Compliance,
New York
M +1-646-227-4345
E johnnyzhang@guidehouse.com

James Robertson
Managing Consultant
Global Investigations & Compliance,
London
M +44 (0) 20 7661 0635
E jamesrobertson@guidehouse.com

linkedin.com/company/guidehouse

twitter.com/guidehouse

guidehouse.com

About Guidehouse

Guidehouse is a leading global provider of consulting services to the public and commercial markets with broad capabilities in management, technology, and risk consulting. We help clients address their toughest challenges with a focus on markets and clients facing transformational change, technology-driven innovation and significant regulatory pressure. Across a range of advisory, consulting, outsourcing, and technology/analytics services, we help clients create scalable, innovative solutions that prepare them for future growth and success. Headquartered in Washington DC, the company has more than 7,000 professionals in more than 50 locations. Guidehouse is a Veritas Capital portfolio company, led by seasoned professionals with proven and diverse expertise in traditional and emerging technologies, markets, and agenda-setting issues driving national and global economies. For more information, please visit: www.guidehouse.com.

Guidehouse