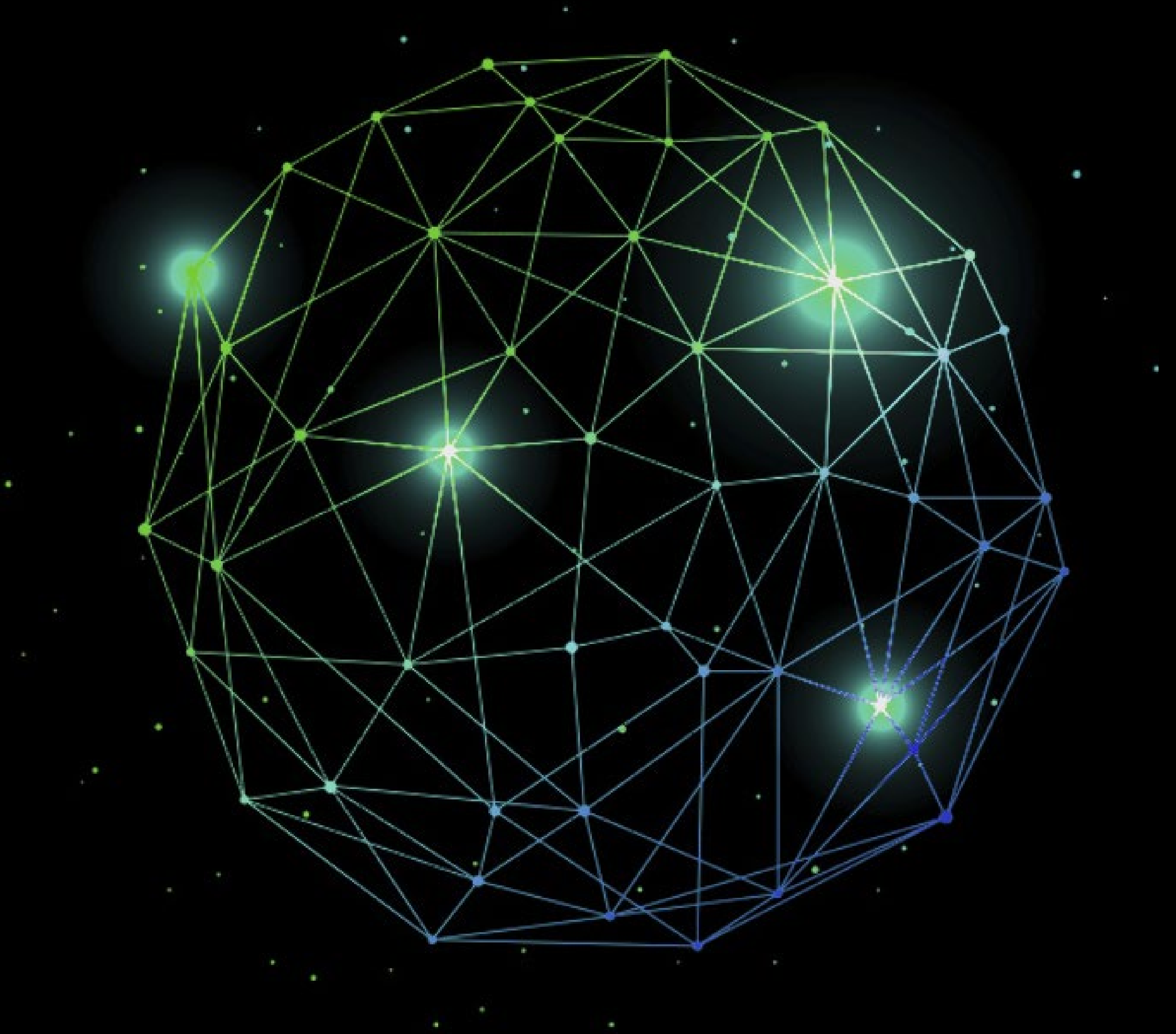


Optimizing rule-based transaction monitoring systems

MAKING AN IMPACT THAT MATTERS since 1845

Introduction

A computer-assisted transaction monitoring system is a key component of internal control environment related to fight against money laundering, terrorist financing, sanction evasion and fraud.

In recent years the matter of enhancement of transaction monitoring systems by the implementation of intelligent components such as machine learning algorithms has gained more relevance than ever. The idea behind using more elaborate systems is the advancement of rather inflexible rule-based systems not only helping financial institutions decreasing costs of false positives as well as false negatives, but also to detect schemes that were not discovered beforehand as there is no typology paper addressing them increasing the overall effectiveness of the system. In our paper we use the terms "machine learning" and "AI" interchangeably.

There are three types of transaction monitoring systems:

- Rules-based systems
- AI-assisted rules-based systems
- AI-based systems

As the acceptance of AI-based systems by regulators is still rather hesitant and artificial intelligence is not too established in the anti-financial crime domain in Germany quite yet, not all financial institutions feel comfortable with the thought of completely replacing their rule-based system with an AI-based one. Thus, this whitepaper is largely dedicated to the second type of monitoring systems, where machine learning works in parallel with (not replacing) core rule-based systems, and assists (but again does not replace) the human component in critical internal control processes.

Overall, AI and machine learning could have multiple uses throughout the entire AML workflow, reaching from the automated generation of investigation documentation all the way through the identification of new money-laundering typologies. As our last publication focussed on using machine learning for evaluating /scoring transaction monitoring alerts[1], this whitepaper further sheds light on an operational approach to perform threshold optimization.

The research and experiments that the authors have conducted had the purpose to enhance the efficiency as well as the effectiveness of transaction monitoring systems whilst retaining existing quality assurance procedures pertinent to a rule-based system. This paper aims at describing in how far a global threshold tuning approach enhanced by machine learning algorithms can be explained on a theoretical level aiming at achieving a larger understanding of what the near future of transaction monitoring systems might look like.

[1] "Künstliche Intelligenz in der Geldwäschebekämpfung", Deloitte, 2023, accessible: Deloitte-Künstliche-Intelligenz-Geldwaeschebekaempfung.pdf

Regulatory framework of transaction monitoring

German anti-money laundering and counter-terrorism financing laws and regulations define a standard that urges financial institutions to implement a transaction monitoring system which scans every single transaction for whether there is any evidence for potential engagement in money laundering, terrorism financing or other criminal offences.

This data processing system must be able to monitor for suspicious behaviour predefined by anti-financial crime experts based on the risk analysis of the financial institution concerning single transactions as well as overall transaction behaviour. Alerts are thereby generated as soon as certain currently mainly rule-based indications, scenarios and parameters are triggered.

According to German legislation, transaction monitoring systems of financial institutions must cover all relevant typologies which are published by the German Financial Intelligence Unit (FIU) as well as other publicly available information on money laundering and terrorism financing schemes. Hence, financial institutions have established teams that are responsible for the regular adjustment and expansion of rule sets closely monitoring and reviewing current trends and publishments within the field of anti-financial crime to meet regulations. However, rules always must be customized to the business strategy, institutional structure as well as risk appetite for them to be as effective as possible. The derivation process must be documented to ensure that the process is coherent and audit-proof.

Overall, the audit-proof documentation of processes and procedures as well as the derivation of rules and thresholds is of utmost importance for financial institutions as a lack of such can be interpreted as not meeting standards and regulations by external reviewers which might cause the incurrence of fines that have to be paid. This is also one of the reasons why the usage of machine learning algorithms to define thresholds is not yet a standard procedure within the German financial industry, as per laws and regulations, there is a prohibition of so-called blackbox technology, meaning that the regulator always must be enabled to understand processes behind the implemented transaction monitoring system. This is achievable, but calls for thoroughly developed systems that are highly transparent and have a sound explainability approach.

In addition to developing new rules, the set of rules in place also needs to be reviewed. This happens either event driven because of adjustments of internal risk policies and appetite within the financial institution or based on a periodic review testing the effectiveness as well as the efficiency of the current transaction monitoring system. The latter is usually executed on a yearly basis and concerns the entire system in place. These processes are rather time intensive and offer potential for an increase of efficiency when assessed properly.

Traditional threshold tuning

As the above-described rule-based system is rather inflexible in nature and does not allow for automatic adaption of new thresholds, there is still a need for regular manual reviews to ensure effectiveness and increase efficiency of the system in place. There are two different testing methods that are conducted to understand whether there is a need for changing the current threshold.

The first test is called below-the-line (BTL) test and is used to understand whether suspicious activity can be detected below the current threshold indicating for the system in place not to be effective. Here, a certain factor x that is derived by statistic models is subtracted from the current threshold in place leading to the occurrence of more alerts. These are then manually reviewed by analysts to understand whether there are cases where the filing of a Suspicious Activity Report (SAR) would be necessary. If so, this might lead to a decrease of the initial threshold to design the system more effectively and ensure that there are less suspicious activities that remain unseen. However, this increases the costs of the overall transaction monitoring, as more analysts are needed to conduct case assessments for the increased number of alerts. Therefore, it is important to ensure an objective testing process to prevent subjective goals such as keeping costs low from influencing the judgement of additional alerts generated by the BTL test.

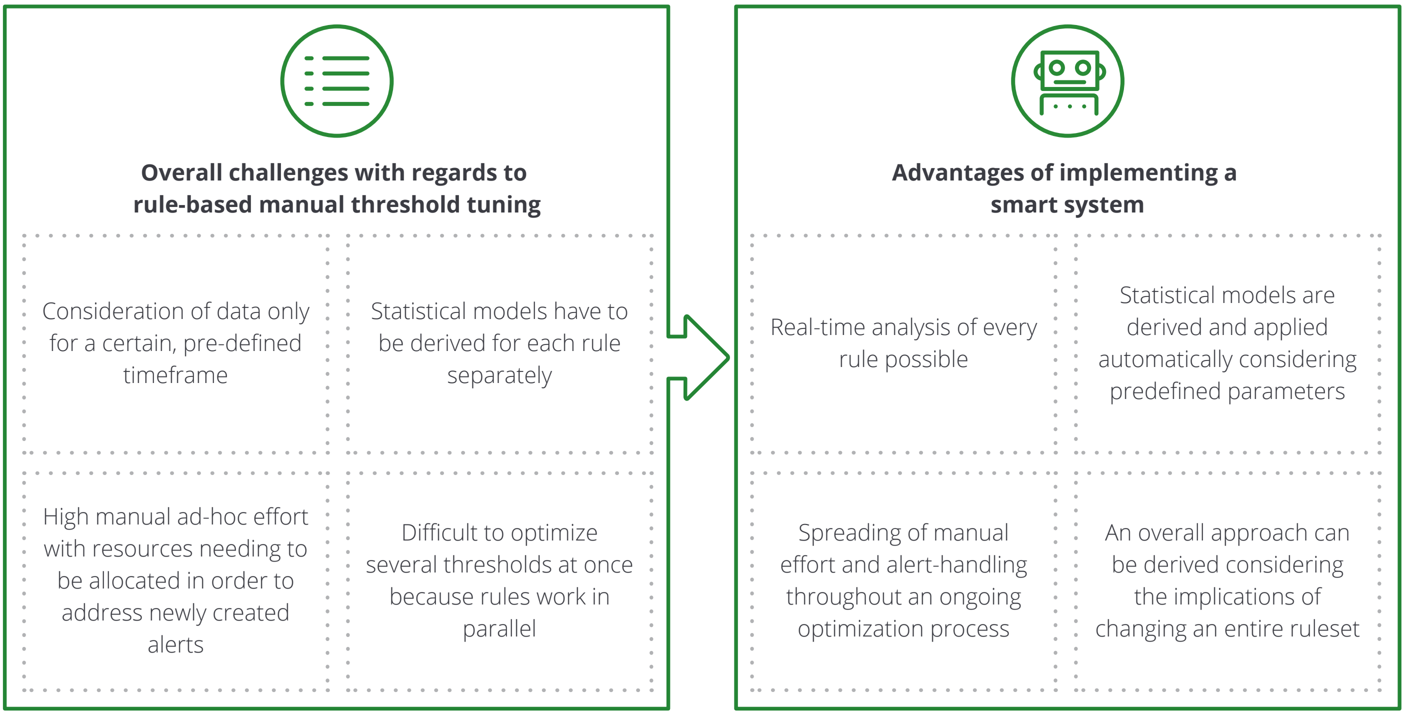
The second test is called above-the-line (ATL) test and describes the process of testing whether the increase of a threshold leads to overseeing suspicious activity or rather solely improves the current alert-to-SAR ratio by decreasing the number of false positives. The testing itself is done by adding a certain factor x to the threshold in place and analysing the alerts generated by the newly set parameter for the rule to understand the effects of the change made. However, when increasing a threshold to claim efficiency gains, the risk of overseeing suspicious activity simultaneously increases. In case failing to detect such suspicious activity can be interpreted as being systematic as it happens regularly, regulators may claim ineffectiveness of the transaction monitoring system in place which can form the basis for fines. Hence, the derivation of a correct conclusion from ATL-testing is rather difficult, especially considering that the costs of false negatives are, in contrast to false positives, difficult to quantify as there are no predefined fines that are to be paid for certain haziness of a system. Hence, basing decisions on a loss matrix is not possible and the process usually is rather subjective and informal. In addition to this, when conducting regular reviews, only a certain, pre-defined timeframe of transactions is considered which makes the system prone to error.

The complexity of the testing processes is increased even further when considering that the single rules should not be adjusted isolated but rather must be understood as being a part of a larger system where the adjustment of one rule might have an impact on the effectiveness of another rule as they might both target the same typology. Therefore, the following chapter describes how these challenges might be addressed by implementing a smart system, allowing for a more integral process.

A framework for enhancing rule-based transaction monitoring systems with AI/ML

One approach in facing the above-mentioned challenges given by threshold tuning within rule-based transaction monitoring systems that is gaining relevance is the introduction of artificial intelligence components. When done correctly, this offers the chance to design a mathematical approach to rule optimization that can be documented and reasoned properly to be compliant with current regulations. Advantages of using an AI approach are described in the following graphic. However, the usage of artificial intelligence should be carefully considered as there are also some prerequisites and challenges that have to be addressed. Hence, this chapter will give insight on what to consider before introducing AI-based systems for a successful implementation. It also sheds light on how AI can improve single rule threshold tuning to lay the foundation for what is possible on a global all-encompassing ruleset.

Fig. 1 – Added value by AI-enhancements of a rule-based system



Overall challenges with regards to rule-based manual threshold tuning

Consideration of data only for a certain, pre-defined timeframe

Statistical models have to be derived for each rule separately

High manual ad-hoc effort with resources needing to be allocated in order to address newly created alerts

Difficult to optimize several thresholds at once because rules work in parallel

Advantages of implementing a smart system

Real-time analysis of every rule possible

Statistical models are derived and applied automatically considering predefined parameters

Spreading of manual effort and alert-handling throughout an ongoing optimization process

An overall approach can be derived considering the implications of changing an entire ruleset

Obstacles to AI/ML in transaction monitoring

Diving deeper into the derivation of a smart threshold optimization model, there are several new challenges that come with the introduction. They not only slow down the replacement of rule-based systems by AI-based systems but also a wider use of AI/ML algorithms for optimizing legacy rule-based systems.

Lack of interpretability of machine learning and AI outputs is often cited as the main obstacle for the use in highly regulated environments, such as AML in banking. However, interpretability of is neither the only nor the hardest of problems related to ML/AI within the anti-money laundering domain. In fact, there are methods to mitigate intrinsic lack of interpretability of many models. Instead, there is a multitude of other hard to solve issues. Let us recall the most relevant ones briefly.

Highly unbalanced datasets

Very few countries ensure that suspicious transaction details are shared across the banking system. Normally, the bank only gathers information about transactions which it has deemed suspicious itself and issued SARs about to the respective authority. The number of SARs is thereby usually limited, particularly for a specific risk topology. In training ML models on known suspicious transactions, one faces what in ML-jargon is called an unbalanced dataset. The proportion of good transactions to bad ones is rather large, which in the context of ML leads to an effect called over-fitting, i.e. the machine learns to react to transaction parameters that are not representative of the risk. This is logical as risk patterns may not be visible from the few transactions that resulted in SAR. Instead, the ML model hallucinates.

Example: "The herd of mules"

Another issue that has to be faced when implementing a machine learning component relates to the classification of suspicious activities. To illustrate this issue, the following section uses the example of animals to describe the issue further.

A classifier model distinguishing cats from dogs, is trained by seeing pictures of cats and dogs. One has to obtain enough labelled pictures. Since cats and dogs do not interbreed, there should be no confusion where the cat is and where is the dog. One could possibly perform a genetic test to resolve an uncertainty.

No such clarity exists between animals that interbreed, such as horses and donkeys. There could be an animal with characteristics of both, such as a mule. Class of suspicious transactions does not have clear boundaries and no genetic test is available to supply an accurately labelled set of suspicious transactions. In mathematical

jargon, such sets are called "fuzzy". Some transactions are "horses", some "donkeys", but many are "mules". There is a field of mathematics dealing with such mules, but it makes already complex ML models just more complex, called fuzzy mathematics.

This problem is directly translatable to a rule-based transaction monitoring system in describing features of suspicious activities. Often, single parameters do not indicate with certainty whether a transaction is suspicious, but in combination with other parameters or throughout the analysis process a decision is made. Hence, some parameters can be understood as indicators and their accuracy can be improved over time. However, as this issue needs somewhat more attention, there will be another whitepaper enhancing this series focussing on fuzzy mathematics in the space of transaction monitoring released at a later point of time.

Accuracy problems with input data

Not only two main classes have unclear boundaries, but parameters describing transactions have intrinsic accuracy issues. These fall into two groups:

- First, there are epistemological problems. These are problems with measurement. Transaction amounts are measured accurately, but a net worth of an individual is probably not. If a database lists EUR 5 million as a client’s net worth, taken from a KYC self-declaration, it informs us that the costumer's worth is unlikely to be EUR 100,000 or EUR 50 million. But it could easily be EUR 5,000,001 or 5,100,000.

- Second, there are ontological problems, which concern parameters that are intrinsically fuzzy in nature. "High risk countries" is an example of such a parameter. Even if the determination of “high” is supported by a score, the process of selecting said score and scoring is typically judgmental in nature. Indeed, behind such a risk score there may be a probabilistic model that links frequency of negative events with a country.

This accuracy problem further caters into the complexity of a machine learning system used for transaction monitoring. As the complexity of fuzzy mathematics and the solution to this challenge exceed the dimensions of this current whitepaper, this topic will be addressed at a later point in a separate publication.

Inestimable risk of missing suspicious transaction

Any optimization problem involves finding a minimum or maximum of merit function for an allowed range of input variables. In the AML domain, a merit function would be cost of risk. Take a single threshold rule, such as monthly amount of cash deposit. It is intuitively clear that increasing the threshold increases risk of missing suspicious transactions and decreases the number of false positives, and, hence, makes rule more efficient but less effective. Decreasing the threshold will achieve the opposite, making it more effective, but less efficient. It seems that there should be an optimal balance between effectiveness and efficiency or, in other words, a global minimum for the cost of risk function. Determining the cost of missing a suspicions transaction (false negative) is the hardest problem, preventing applying traditional linear programming methods for deciding on an optimal set of thresholds. This challenge will be the one this whitepaper will focus on to achieve a possible framework for a global rule optimization approach, which can later on be enriched by solutions to the above-mentioned issues.

Key obstacles to AI/ML in transaction monitoring is the handling of highly unbalanced datasets, the inestimable risk of missing suspicious transactions as well as accuracy problems with input data.

AI-assisted BTL testing for manual threshold optimization

As per our project experience, most banking institutions do not yet use machine learning and other quantitative methods to optimize their AML transaction monitoring systems given the multiple issues we just listed:

- Unbalanced and fuzzy classes.
- Inaccurate values of input parameters feeding into detection models.
- Cost of false negatives cannot be quantified.

One of the reasons why AI/ML in transaction monitoring face credibility issues with regulators is because quantitative methods pretend to use accurate and balanced inputs, where such do not exist. Instead, many banks chose to retain a qualitative approached based on a judgmental process of optimizing the system with the help of ATL/BTL tests.

In a first step, even such a qualitative approach could be improved by automating BTL testing using machine learning methods. Below we describe our proposed methodology and the proof-of-concept that has been developed and deployed in cooperation with one of our technology vendors.

Theory of optimizing threshold values for single rules

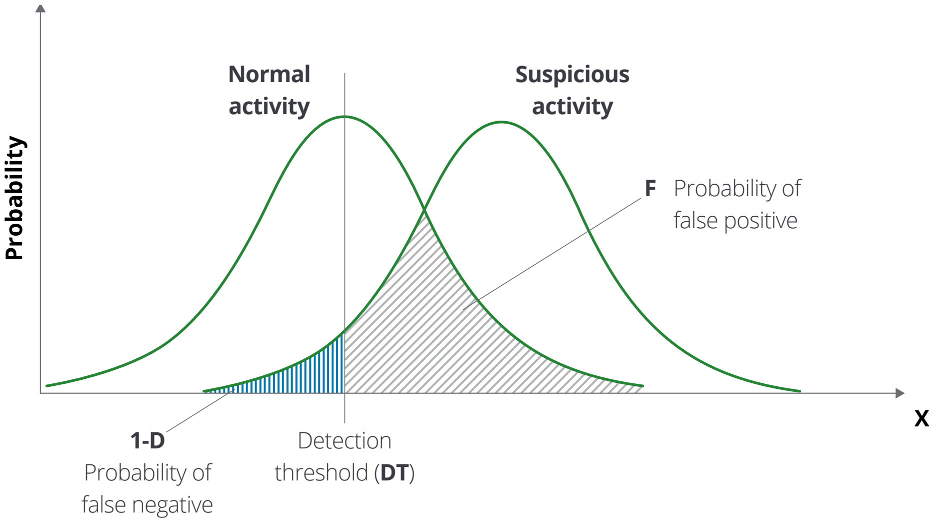
A productive way of looking at optimizing detection rates is through a probabilistic lens. The chart below illustrates this approach for a single parameter of a transaction. In real life, there are tens to hundreds of parameters feeding a transaction monitoring system. Through so-called feature engineering, a machine learning model used in transaction monitoring may be provided with thousands and more parameters. The basic principles would remain the same also in a multi-dimensional parameter space.

The two bell-shaped overlapping curves illustrate probability density functions for parameter as related to normal or suspicious activity. The graphs will overlap, which means that, based on prior experience, a single value of a parameter could correspond to both normal or suspicious activity. If they do not overlap for at least one of many parameters (a “miracle” parameter), the task is optimizing the threshold is trivial. But this is never the case.

If nothing is known of the consequences of making a detection error, the threshold is best set in such manner so the shaded areas (integrals of the probability density function) to the left and to the right of the threshold are equal from a value perspective.

If one could make an inference regarding the cost of making wrong conclusion about a transaction it would be possible to move the threshold to the left or right proportionate to the relationship between high cost of false negatives and comparably low cost of false positives. The formula for the optimal threshold minimizing the overall risk in such manner is presented above. But as explained in the preceding chapter, cost of false negative is hard to estimate.

Fig. 2 – Graphical description of probabilities of false positives and negatives



Learning from analysts' experience
One of the approaches is to re-define the problem. There could be two tasks that the transaction monitoring system should solve:

- Improving the quality of detection comparative to solely rule-based systems with manual resolution. In technical terms such system would focus on reducing false negatives (FN) or suspicious transactions that go undetected.

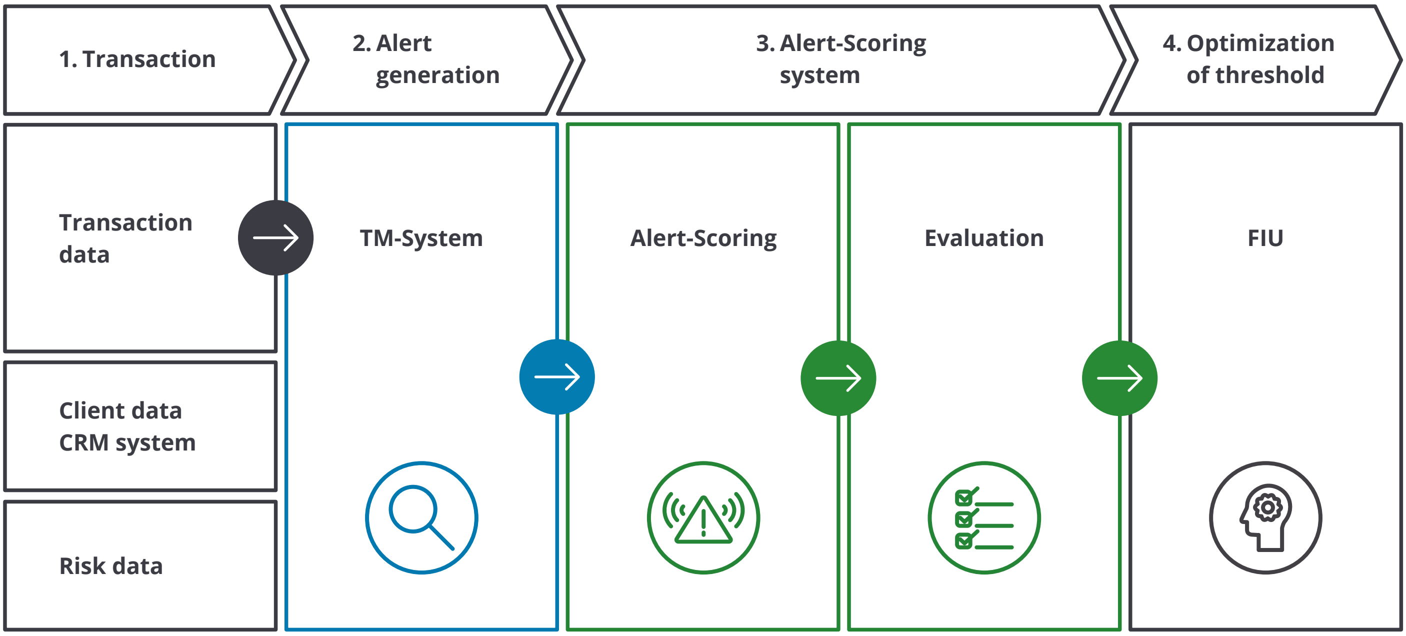
- Improving efficiency of the system and reducing human effort that is spent to false positives (FP) or transactions that are initially identified as questionable but later resolved to be good. If the system is more efficient, the human effort spent on resolving FP could be better directed towards problematic transactions, in such way improving FN ratio. Effectiveness and efficiency are related.

When the problem is being re-defined as reducing FP, the training set is no longer unbalanced since there are many good transactions to learn from. Some of the transactions that triggered an alert resulted in an investigative case and a lengthy review process. As per our experience, the ML model could be trained with this information so transactions that do not result in a case investigation are routinely suppressed. With such an approach, the AI learns to re-perform investigator's activity, with faster response time, at lower cost and with greater consistency. Learning to do such task, the machine could also help building and maintaining up-to-date whitelists of transactions and counterparts.

The AI/ML could also make inferences on the risk level of a transaction by observing time spent by a human investigator on clearing the alert. If a certain type of alerts is routinely supressed without further investigation, this might be indicative of a false positive.

Most of the human investigation occurs within the system and there are digital footprints of the activity available for learning. We have developed a machine learning model that has been observing time spent on alert review. Timestamps available in most systems allow for such analysis at no cost. Analysis of investigator behaviour could range from such simple timestamp analysis to complex process mining. In any case, it may be able to produce a training set of transactions that has been routinely dismissed by investigators.

Fig. 3 – Example of Deloitte developed Proof of Concept



Description of Proof-of-Concept software

Deloitte has developed a proof-of-concept (PoC) software approach for such described threshold optimization. At the core of the PoC is a machine learning model for transaction monitoring alert triage. For training purposes, it uses investigation cases and SARs. As output, the machine learning model provides alert scores, which are further used to put alerts into three different classes: Red, Yellow, and Green.

The model is being used to score alerts below and above the threshold. Scoring above-the-threshold is trivial. Red (i.e., alerts above certain risk score) is therefore especially used as a proxy for suspicious activity below the threshold. In such manner, the system also allows to perform BTL testing on a continuous basis.

When the machine learning model for alert triage is trained on SARs, the focus is on Alert-to-SAR (ATSAR) and when it is trained on investigation cases, it uses Alert-to-case (ATC) as a representative metric. The right metric is a question of judgment and the available training set. The underlying principle of the system is that true positives and alerts scored Red exhibit similar behaviour. In other words, ATSAR and ATC ratios in population above the threshold and in the population being tested (below the threshold) are consistent.

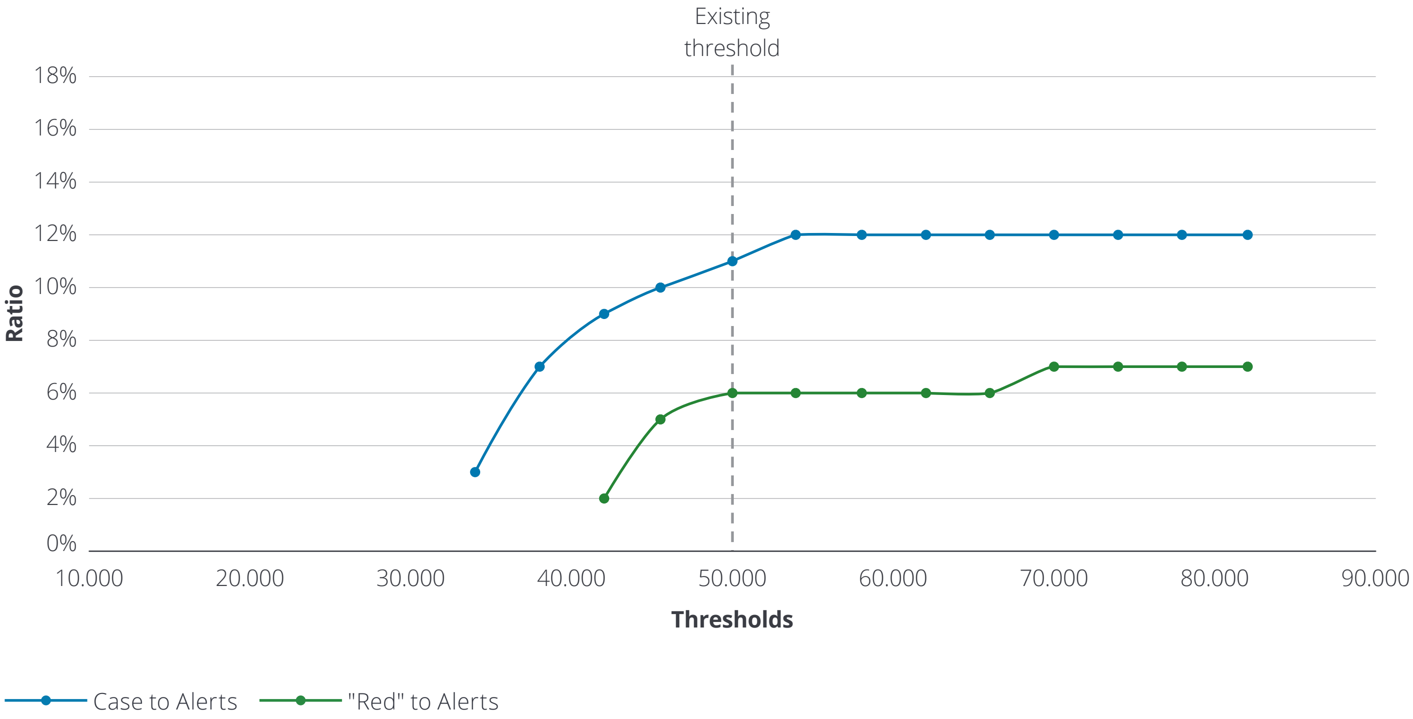
The POC is using a scenario modelling function in the software package used to test each threshold (90%, 80%, etc.for Alert-to-Case and Alert-to SAR efficiency.The POC has been tested on a synthetic transaction data and has demonstrated that it successfully serves its objective.

AI-assisted BTL testing for threshold optimization improves the traditional approach by automating BTL testing and allowing for optimization of detection rates through a probabilistic perspective.

Single threshold optimization process
An illustrative example of ATL and BTL test output (in blue and green respectively) is presented in the chart on Fig. 4. The objectives of the analysis include achieving desired rule efficiency and consistency across the rules, which individually or jointly cover AML risk topologies. Judgement is applied to retain inefficient rules or thresholds if these are expected to address certain topologies.

An AI assisted system may generate an additional metric/curve – alerts to Red (ATRED), a ratio of alerts scored Red to total, which is used as the proxy for ATC or ASAR for BTL area (depending on the labelling of the training set).

Fig. 4 – Illustrative example of the ATL/BTL dashboard



Global AI-enabled optimization of TM rule-based system

Optimizing financial cost of AML risk

Ultimate objective of threshold optimization is minimising the cost of AML risk. In mathematical language it could be re-phrased that the optimization should have financial cost of risk as the merit function. The best transaction monitoring system is one that ensures the lowest cost of risk to the bank. With a single threshold, the financial cost of AML risk could be quantified as following:

$$\mathrm{R}\ (\mathrm{x}) = pC_{10}(1 - D(x)) + qC_{01}\ \mathrm{F}(\mathrm{x})$$

For a system with multiple rules and thresholds, there a will be a system of multiple equations. All elements of the equation except F (x) and D (x) are constants and must be measured. The algorithm described in preceding chapter allows to analyse D(x) and F(x) and visualise behaviour of detection error rates depending on threshold.

The cost element that is most difficult to quantify is C_{10}, the cost to the bank arising from regulatory penalties, reputational damage etc. An approach to address quantification of this risk is presented further in this paper. It must be decided whether the cost of errors is uniform across all rules or is set specific to each rule.

Tab. 1 – Description of equation inputs

Parameters	Description
p	A priory probability of negatives (no suspicious activity)
C_{10}	Cost of false negative (missing suspicious activity). Cost could be defined as financial effect of regulatory sanctions and reputational damage
D	Probability of correct detection of suspicious alerts
q	A priory probability of positives (suspicious activity)
C_{01}	Cost of false positive is time cost to the bank of suppressing false alerts and processing the false positive cases
F	Probability of false positives

Estimating the cost of false positives

Cost of false positives could be obtained by measuring time cost expended by the bank on supressing alerts and investigating cases that do not result in SAR. There are two techniques typically used in this process:

- Process mining;
- Activity-based costing (ABC).

Process mining includes analysing information contained in computer system, such as time stamps or activity logs, to measure time spent by analysts in processing alerts, cases and SARs. To assist with process mining one could use commercially available packages. Time could also be measured manually by collecting manually filled timesheets on a sample basis.

Time needed to investigate false positives is not the only enterprise resource incurred in managing AML control risks. Activity-based costing (ABC) methodology offers a systematic approach for allocating costs of resources, other than direct variable labour. Overheads that have been allocated in order to correctly price AML risk costs include computer systems, specialised and general-purpose, senior management time, office facilities. Some banks instead of performing ABC exercise rely on simple benchmarks such as standard overheads or doubling direct compensation cost.

Estimating the cost of false negatives

With regards to false negatives, the estimation is rather difficult. While this is the component of transaction monitoring models that regulators focus on in order to ensure there is no systematic error, a financial institution would have to review every single transaction in order to ensure that there are no false negatives at all. From a financial viewpoint, this approach does not make sense in order for a bank to stay competitive.

An intuitive approach to evaluating the cost of missing suspicious transaction is through making inferences between single incidents of false negatives and systematic errors to then evaluate the risk of regulatory fines and reputational loss associated with the identified errors. However, we cannot identify or recommend a robust methodology that soundly addresses the above-described challenges. Instead, we suggest an alternative approach that could be implemented using ML technology.

Factoring in risk appetite

One way to describe an institution with “high AML risk appetite” is as one that offers products and accepts customers with high incidence (q) of AML risk events. The regulator would be content with high risk provided it is mitigated by stricter AML controls. There are rewards incidental to servicing high risk segments as well as higher risk costs, including costs of compliance and costs associated with risk events.

Another way to describe the high risk appetite is by reference to risk cost optimization strategy chosen for a given risk profile. Any rational player will minimise risk costs by optimizing controls, including selection of rules and thresholds. But risk costs associated with risk events are not known in advance of these events – they may be deferred and not even be known well past the event. Management has to exercise their judgement particularly regarding C_{10}. Low risk appetite is an appropriate description for a bank that conservatively estimates C_{10}, and vice versa. This concept of a risk appetite is helpful in explaining our approach to evaluating cost of risk.

Assume that the bank believes the risk profile of its customer and product base is appropriate. Assume that the bank believes its risk appetite is appropriate and current rule thresholds are close to optimal. If incidence of risk events is unchanged, for a rule that addresses a given risk event topology, the bank would only change threshold if the relationship C_{10} /C_{01} changes. Assuming no need to change the threshold, we could infer the bank judgement regarding C_{10}. Imagine a cumulative monetary threshold of EUR 50,000 for a monthly cash deposit rule. If the bank regards it as an appropriate than, in a general case, it should regard as comparable any other threshold in close vicinity to the current one, such as EUR 50,001 or EUR 50,010. What logically follows is accepting the balance between cost of risk and cost of compliance around this threshold. Without introducing functional analysis, this could be illustrated with a simple arithmetical example.

Quantifying the cost of false positive

Making an incremental increase of the threshold we should expect an incremental decrease in false positives ΔF and incremental increase in false negatives ΔD. In our fictitious example, increasing a threshold EUR 50,000 to EUR 50,010 we will observe ΔF and ΔD. If we believe that the current level is appropriate than incremental trade-offs around the existing threshold are also appropriate and knowing C_{01} allows us to infer

$$C_{10}:$$

$$C_{10} = (\frac{\Delta F}{\Delta D})C_{01}$$

When the bank analyses results of ATL and BTL tests and is visually studying curves, what it does – it performs qualitative analysis of F(x) and D(x) functions in the vicinity of the existing threshold. We propose to do the same using quantitative methods. The approach assumes F(x) and D(x) functions are not piecewise linear and such an assumption typically agrees to our intuition regarding most of money laundering topologies. Imagine a cumulative monetary threshold of EUR 50,000 for a cash deposit rule. We could expect that probability of someone laundering cash in the amount of EUR 50,010 is very similar to laundering in the amount of EUR 50,000. (Indeed, if the threshold applied by the bank becomes known and criminals were to consciously avoid depositing in monthly amounts exceeding EUR 50,000, D(x) would become piecewise linear.) Absent such condition D(x) is normally smooth and monotonous, at least at the interval around the threshold. ATL allows to model behaviours of both functions very well above the threshold. Algorithm described in the preceding chapter allows to perform the same also for values below the threshold.

F(x) and D(x) are probably non-linear. We could approximate them by linear functions for short intervals around existing thresholds or, to make model more accurate (and more computationally complex) we could fit them with non-linear functions.

The method has obvious weaknesses. It assumes some thresholds are close to optimal allowing inferences regarding implied cost of false negatives. Making judgements which rules should serve as benchmarks or any rules should be grouped for similar cost of risk would require expert judgement. It is applied alongside machine learning model – AI is not expected to replace human expertise, at least, now.

AI-based global optimization of thresholds

Once all the constants are known, the problem of finding optimal thresholds becomes tractable. In mathematical notation the task of determining the optimal threshold could be written as follows:

$$\text{Arg min}\sum_{k=1}^{n} \text{R(x)} = pC_{10}(1 - D(x)) + qC_{01}\ \text{F(x)}$$

where n is the number of rules

This means finding a set of thresholds for an existing rule-based system that minimise the overall cost of risk to the bank, finding the right balance between effectiveness and efficiency. The optimization is being solved using well known linear programming algorithms.

Given the complexity and interdependence of the existing rulesets and the amount of data processed, the usage of artificial intelligence in form of a ready-to-use engine makes sense in this case in order to accelerate the process significantly.

AI-enabled optimization of TM rules presents a comprehensive approach to balance effectiveness and efficiency by setting the cost of false negatives relative to false positives and harnessing optimization algorithms to find ideal rule values.

Conclusion

The need for transaction monitoring within financial institutions is inevitable. From a regulatory viewpoint, there are certain conditions that have to be met overall in order to be and stay compliant, also suggesting to have an IT solution in order to tackle the approach of screening and monitoring all occurring transactions. As financial institutions are legally obliged to cover certain typologies within their systems, these are often rather inefficient and partially ineffective, especially with regards to discovering new patterns. Therefore, implementation of an artificial intelligence-based approach could help financial institutions lower their false positive alerts increasing efficiency, whilst simultaneously allowing for pattern recognition, which ultimately results in higher true positive rates and thereby decreases the probability of being fined for strategically overseeing suspicious transactions.

A first step towards implementing such smart system is by using AI to enhance the process of threshold optimization, with the initial rule-based system still remaining in place. However, at least in Germany there is not yet a framework describing how to implement intelligent solutions that further enhance the performance. Therefore, it is important to consider potential pitfalls and deciding upon an implementation strategy before starting the implementation process. One overarching issue that may occur within the application of AI for transaction monitoring are thereby the occurrence of unbalanced datasets, meaning that in comparison to the number of transactions there is a low number of SARs filed. Depending on the number, this might make it difficult to effectively design AI models without over empathising certain factors. Moreover, it has to be taken into account that not all true positives are definable with certainty as there is usually no parameter that directly indicating the occurrence of suspicious behaviour, but they are rather an indication. A third factor that should be considered is the fact that especially with regards to client information, not all input data is accurate to an extend where it can be anticipated that for example the declared salary matches exactly with what a customer earns. This inaccuracy also needs to be addressed within a sound system.

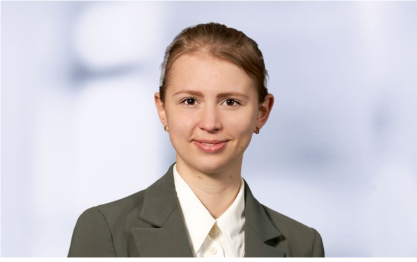
When having defined the parameters around data input, it has to be decided which part of the transaction monitoring system shall be enhanced by the AI component. A first way of doing so is enhancing threshold optimization by alert triage or single threshold optimization using input produced by analysts. However, as the optimization aims at decreasing false positives and ultimately reducing costs, one limiting factor is that the cost of false negatives, meaning overseeing suspicious activity, are not easy to value. Therefore, one approach might be to introduce a global threshold optimization approach where the aim is to keep the number of false negatives stable. In order to achieve this, the cost of false negatives is set into relation to the cost of false positives, allowing for optimization algorithms to find an ideal value for thresholds of different rules at once.

Overall, AI offers great potential for transaction monitoring systems. If introduced properly and planned beforehand, the implementation of smart systems can help reducing costs whilst improving the effectiveness of the tools. However, there is a need for a close collaboration between domain knowledge and technology knowledge in order to achieve a comprehensive system that makes use of the potential whilst also allowing for financial institutions to be compliant with regulations. This paper has therefore given an insight on how this synergy could potentially be achieved.

Contacts

Martin Hirtreiter
Partner | FSI
Tel: +49 69 75695 7059
mhirtreiter@deloitte.de

Dr. Robert Schmuck
Director | FSI
Tel: +49 89 29036 6245
rschmuck@deloitte.de



Janina Uspelkat
Senior Consultant | FSI
Tel: +49 40 32080 5555
juspelkat@deloitte.de

Tim-Lasse Bohm
Senior Consultant | FSI
Tel: +49 40 32080 4114
tbohm@deloitte.de

Igor Rodin
Academic Advisor on AI |
Riga Transport and Communications Institute
Tel +371 29 448703
Igors.rodins@rbs.lv

Deloitte.

This communication contains general information only, and none of Deloitte GmbH Wirtschaftsprüfungsgesellschaft or Deloitte Touche Tohmatsu Limited (DTTL), its global network of member firms or their related entities (collectively, the "Deloitte organization") is, by means of this communication, rendering professional advice or services. Before making any decision or taking any action that may affect your finances or your business, you should consult a qualified professional adviser.

No representations, warranties or undertakings (express or implied) are given as to the accuracy or completeness of the information in this communication, and none of DTTL, its member firms, related entities, employees or agents shall be liable or responsible for any loss or damage whatsoever arising directly or indirectly in connection with any person relying on this communication. DTTL and each of its member firms, and their related entities, are legally separate and independent entities.

Issue 03/2024