WHITE PAPER

APPLICATION OF MACHINE LEARNING IN AML TRANSACTION FILTERING

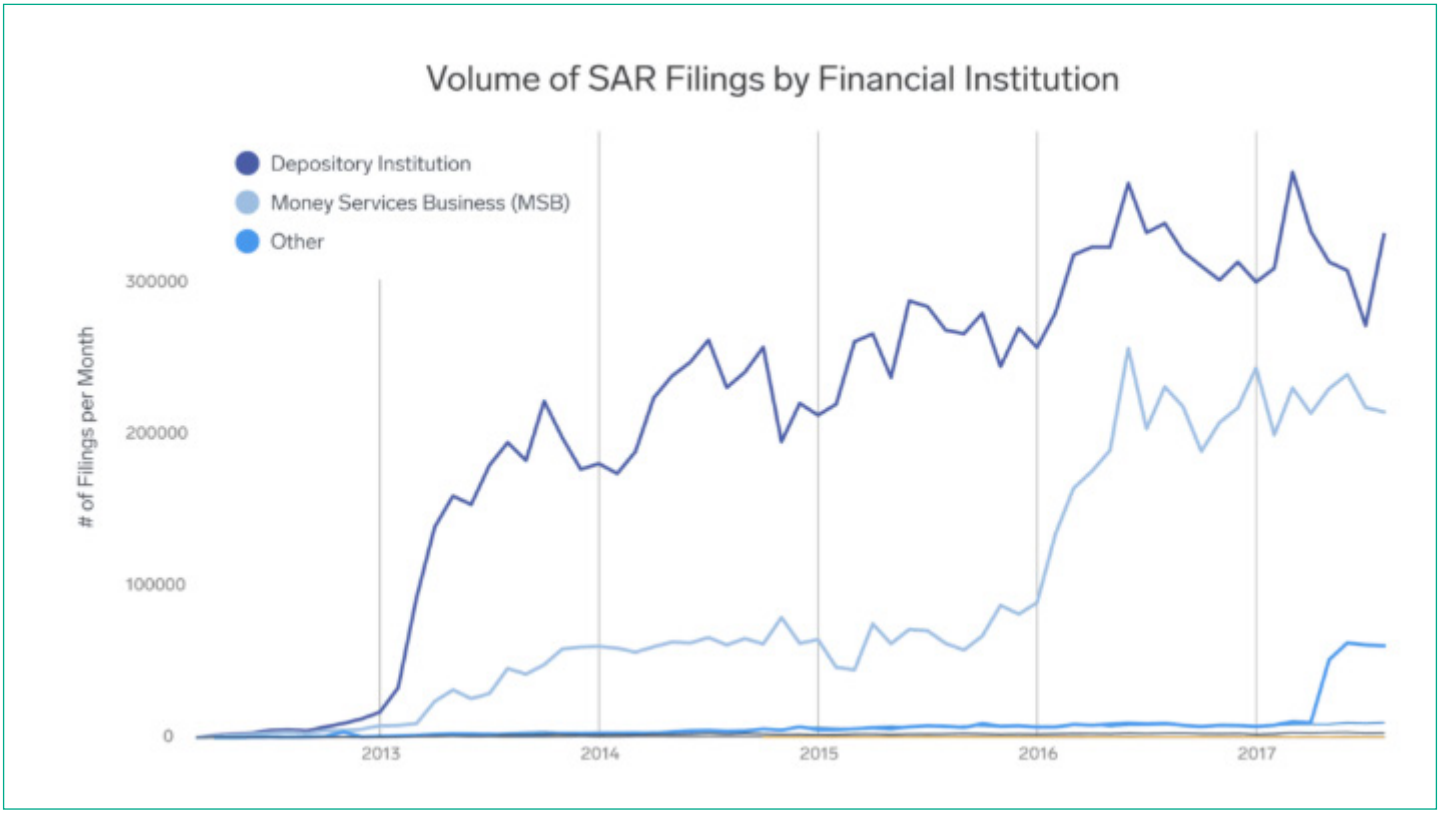
Infosys®
Navigate your next

Overview

The menace of money laundering continues unabated. As per "Basel Anti-Money Laundering (AML) Index 2018"[1] report, money laundering and terrorist financing continue to distort international finances, cripple economies and harm citizens across the globe. According to estimates[2], the amount of money laundered across the world ranges from US$500 billion to US$1 trillion.

Over the past seven years, since it was first computed, the Basel AML Index has consistently shown slow progress among most countries in enhancing their money laundering / terrorist financing risk scores. 64% of the countries in the 2018 ranking have a risk score of 5.0 or more – implying significant risk of money laundering and terrorist financing. Also, 42% of the countries have seen deterioration of their risk scores between 2017 and 2018. The Basel AML Index has shown that there is low level of effective AML/CFT measures enforcement.

Over the past few years, the number of suspicious activity reports (SARs) filing has been increasing in many jurisdictions. As per a Financial Times article[3], in 2018, the number of SARs filed to the United Kingdom National Crime Agency reached record level of ~464,000 – up ~10% from 2017.



Source: *https://www.enigma.com/blog/trend-watching-across-fincens-suspicious-activity-data*

Exhibit 1: *Growing volume of SAR filings by financial institutions (as per FINCEN's data)*

Growing importance of AML transaction filtering

Given the above, no wonder the AML software solution market is expected to grow significantly in the coming years. As per ResearchAndMarkets.com's report[5], the global AML software market was valued at US$898 million in 2018 and is forecast to grow at a CAGR of around 14% to reach US$1.9 billion by 2024.

AML transaction filtering is a key component of any robust AML software solution, and is an important part of the AML checks implemented by a financial institution (FI). It comprises sanctions and blacklist screening and customer profiling. It screens transactions at pre-execution stage to prevent activities in violation of the AML rules.

AML transaction filtering is also a regulatory requirement. FIs need to have adequate checks and filters in place to detect and prevent dirty money from entering the banking system. Not doing so would attract hundreds of millions of dollars in fines for the concerned FI, apart from reputational loss.

Key benefits expected from AML transaction filtering include:
a) ensuring global compliance consistency by screening transactions against global watch-lists b) blocking of transactions in real-time c) compliance costs reduction d) speedy investigations and reporting.

Source: *https://www.alliedmarketresearch.com/anti-money-laundering-software-consumption-market*

Exhibit 2: *AML software market size is growing*

Need for AI/ML adoption in AML transaction filtering

Over the years, regulatory costs due to AML non-compliances have increased exponentially. Regulatory expectations too have increased manifold. This is evident from the quantum of fines being imposed on FIs for non-compliance.

As transaction volumes rise, it is becoming increasingly difficult for FIs to rely on legacy filtering systems and processes that are found wanting. Criminals are finding innovative ways to defeat the checks that are in place in these legacy systems.

Further, according to a report from PricewaterhouseCoopers (PwC)[7], almost 90-95% of the alerts generated are false positive. This not only leads to a huge operational overhead for banks, but they also run the risk of missing genuine alerts as they wade through the alerts list.

To overcome the challenges mentioned above, FIs are looking for new solutions. Artificial intelligence and machine learning (AI/ML) based solution can help FIs in this regard. For example, AI/ML based solution can be leveraged to reduce false positives and improve the quality of the alerts.

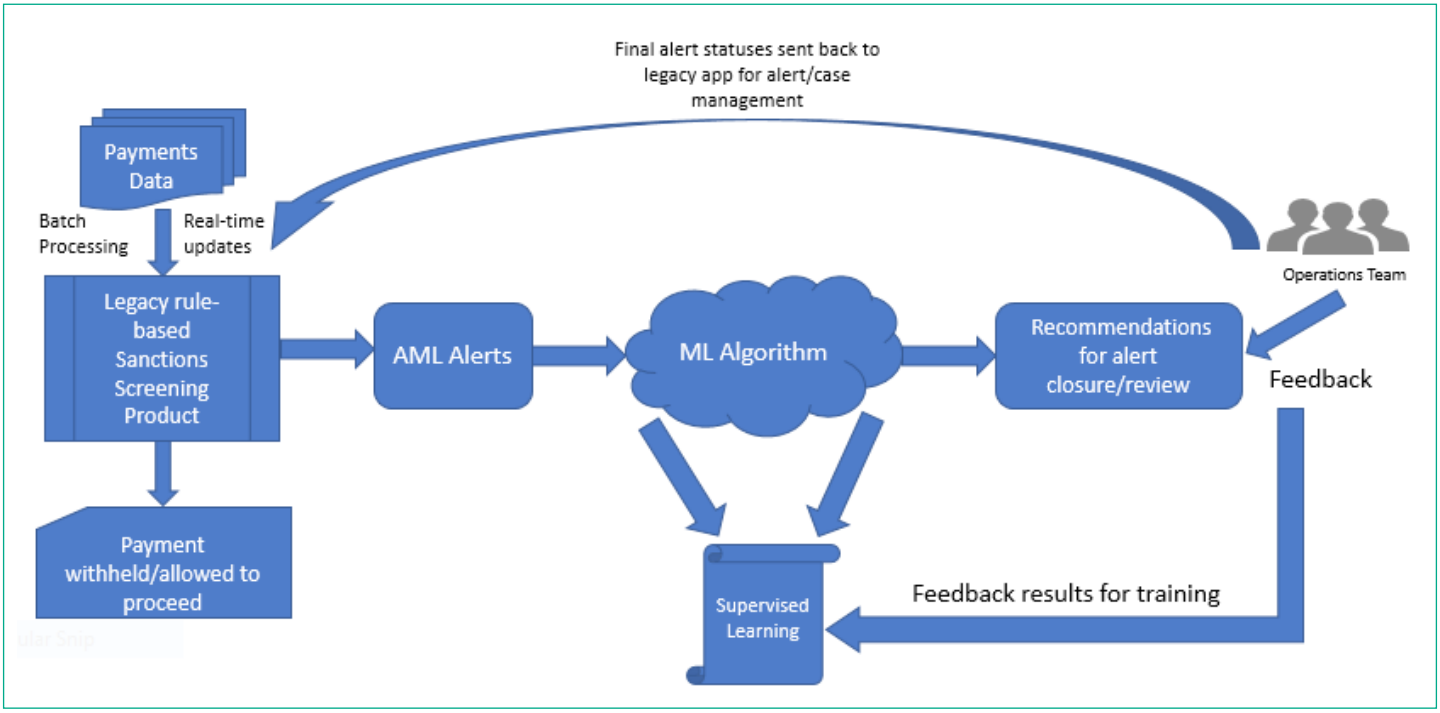


Exhibit 3: *Illustrative ML adoption in AML alerting process*

This PoV takes sanctions screening as an example. However, the recommendations provided can be applied to other AML based transaction filtering aspects as well.

Traditional AML sanctions screening workflow

As part of transaction filtering, FIs are required to perform a sanctions screening to check if any customer is involved in financing based crime. FIs have an obligation to screen payment instructions prior to execution, in order to prevent breaching sanctions, embargoes and other AML measures.

Sanctions screening is the process of reviewing the bank's payments on a real-time basis against a sanctions list to check if the payee or beneficiary is involved in financing of crime or terrorism. If the screening results in a positive hit or a suspicious transaction, the payment can be blocked and investigated further. These sanction lists are procured from major regulatory bodies.

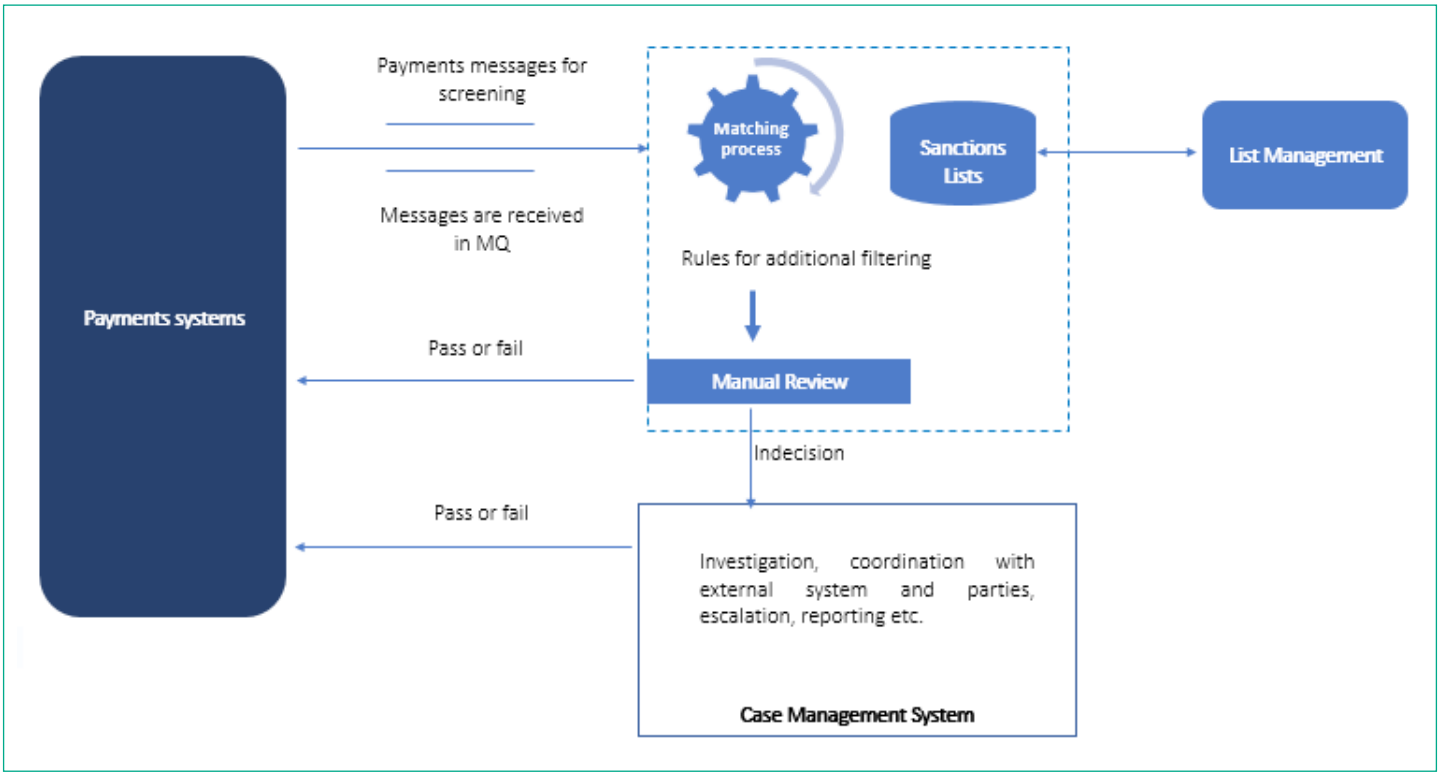


Exhibit 4: *Traditional AML sanctions screening workflow*

Typically, the rules for sanctions screening are based on various text matching algorithms, few examples of which are given below:

1. Restrictive exact match: This generates a positive match when the input data exactly matches the person on the sanctions list. This takes into account possible name juxtaposition.

For example, using this algorithm, all the following names would match "Osama Bin Laden".

- Laden Bin Osama
- Bin Osama Laden
- Osama Laden Bin

2. Fuzzy match: This allows the algorithm to determine similarity between data elements. It detects and evaluates near matches instead of exact matches. The percentage of match can be set by the organization.

For example, using this algorithm, Osama Bin Laden would be matched with

- Osama Been Laden
- Osam Bin Leden

Issues with traditional AML sanctions screening workflow

While there are many such algorithms and name matching techniques which can be used, they all suffer from a common issue: false positives.

False positives occur when a transaction associated with a genuine customer is blocked because of a name or any other type of match. For example, payment to a client living in Kerman, California could be blocked due to a match with the city of Kerman, Iran. They are a significant burden to FIs since they result in significant cost and effort to clear possible violations which are actually genuine. This process delays the payment release. Additionally, the FI has to engage more human investigators which results in further monetary expense.

Over the past few years, there has been manifold increase in the number of sanctions screening alerts. There are multiple reasons for this. Few are outlined below:

1. Payment transactions are increasing over the years.
2. Risk entity entries in the sanction list are increasing with time.
3. The matching processes are sub-optimal due to fuzzy logic thresholds not being set properly.
4. For many types of payment transactions, due to data quality issues, the payment format standard is not recognized. Due to this, field scanning is not being undertaken.

A unique approach to AI/ML based AML sanctions screening

AI/ML based solution can be leveraged for reducing sanctions screening related false positives. There are multiple products in the market that have begun offering AI/ML based sanctions screening solution. However, many FIs that have already invested and drawn-up contracts with traditional products vendors may lack advanced ML characteristics. Yet, such FIs may not be eager to change the product to a new one – considering the risks and cost involved.

An **ML plug-in approach** can be adopted in such cases. Once the ML plug-in has been built, the reference functional architecture could be as below.

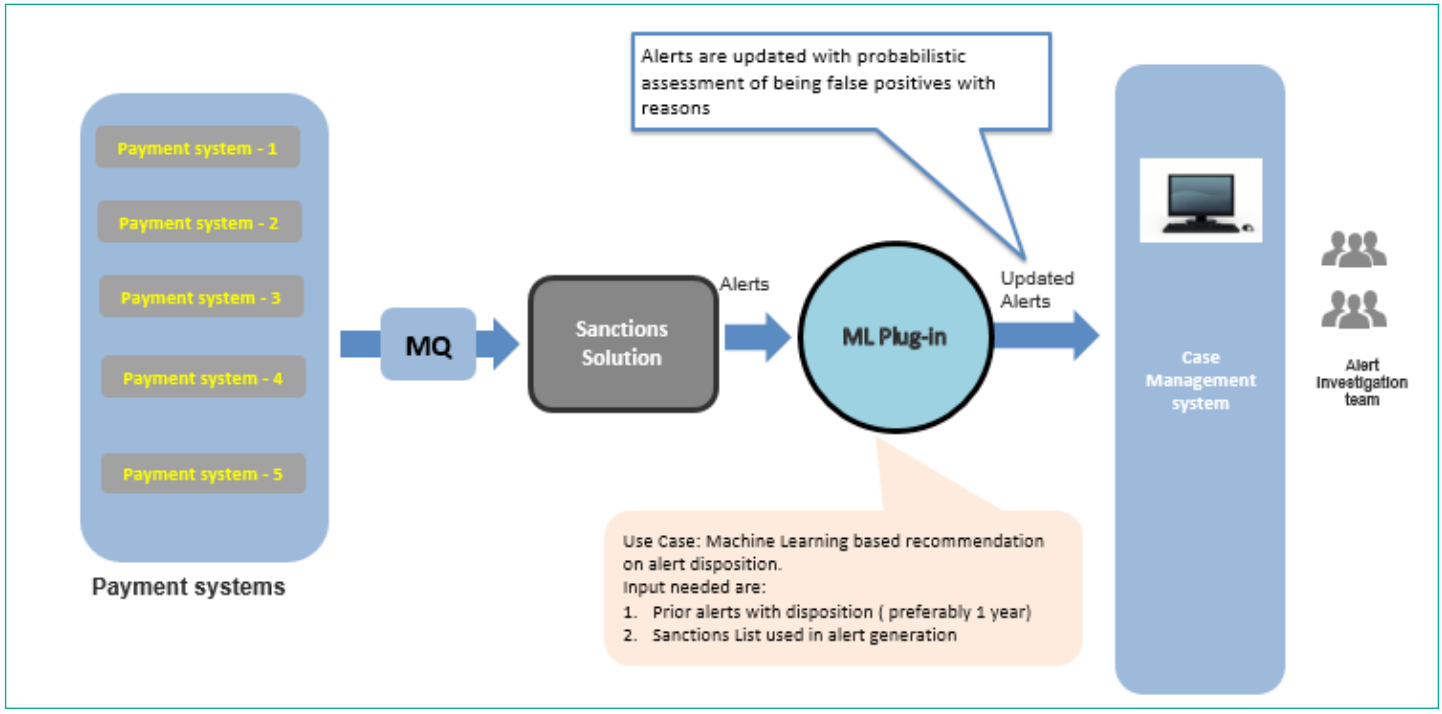


Exhibit 5: *Reference functional architecture with ML plug-in*

While it is extremely difficult to completely eliminate the false positives when taking a balanced risk-based approach to sanctions screening, there are a few ways to reduce the numbers:

1. Duplicate alerts: If the alert for the same person gets generated every time they make or receive payments and it has been determined that the alert is false, then the ML algorithm can be trained so that it gets closed automatically if such an alert is generated again in the future against the same entity, assuming that there are no modifications to the said entity.
2. Data capture in multiple fields: Data can be captured in multiple fields. For example, name can be gathered as title, first name, middle name and last name. This helps to avoid ambiguities. Additionally, the ML algorithm can be taught to match on a combination of multiple fields like Name and Customer Type. This is to avoid false alerts where names could often be similar for individual and corporate customers.
3. Alert data enrichment: Every alert when closed by the operations team, must have enough comments and details about why a particular action was taken. Similarly, every entity must have history of all the past alerts and actions taken on them. This history is useful when analyzing a repetitive false alert. The ML algorithm can take decisions based on past actions.
4. Reason categories for false positives: When an alert is closed as false positive by the operations team, the reason should be clearly stated. The main repetitive reasons for false positives should be identified and a drop-down field should be provided while closing the alert. This will help in identifying root cause and help in trend analysis. The ML algorithm can learn and suggest auto closure of similar alerts in the future.

Implementation approach

For the system to start learning, the main input needed are prior alerts, preferably 1 years' worth of data and the sanctions list. The sanctions SME, the AML operations team and the data scientist with ML expertise need to work together to understand how the alerts are classified as false positives. Additionally, they would also learn the main reasons for false positive alerts. These learnings will be applied to create a supervised ML algorithm based on the techniques above.

These algorithms must be tuned from time to time so that the system "learns" to distinguish correctly between genuine alerts and false positives. Without periodic checking, there is high probability of falling into the other dangerous zone of false negative. With a combination of periodic manual checks, automation of repetitive processes and constant teaching of the system via historical data, FIs should be able to reduce the false positives and improve their sanctions screening process.

Some examples of how the algorithm can be taught are given below.

1. If the false positive alerts are mainly due to partial matches, the matching algorithm will be modified to have a higher match percentage or taught to take into consideration the complete name to understand if the customer is actually part of the sanctions list.

In the above example, the algorithm will be taught to consider all the 3 fields – first, middle and last name to avoid a genuine customer's payment from being withheld.

Sanctions list:

First Name	Middle Name	Last Name
Osama	Bin	Laden

Customer name:

First Name	Middle Name	Last Name
Osama	Bin	Mahmood

2. If the name matches completely with the sanctions list and causes false positives, then the ML algorithm may be trained to consider additional fields to check if this is really a genuine alert.

In this example, the algorithm will be taught that since the customer has most of his transactions based in New Zealand, he should not be considered as high risk leading to withheld payments.

Sanctions list:

First Name	Middle Name	Last Name	Country
Steve	M	Law	Columbia

Customer name:

First Name	Middle Name	Last Name	Country
Steve	M	Law	New Zealand

3. Sometimes the customer name could partially match to the name of a corporation leading to false positives. Once again in this case, the ML algorithm must be trained to consider additional fields to check if the alert is genuine.

Sanctions list:

First Name	Middle Name	Last Name	Customer Type
Omar		Technologies	Corporate

Customer name:

First Name	Middle Name	Last Name	Customer Type
Omar		T	Personal

In the above example, the algorithm will learn that even though there is a partial match, the sanctioned customer is an entity, while the genuine customer is a person.

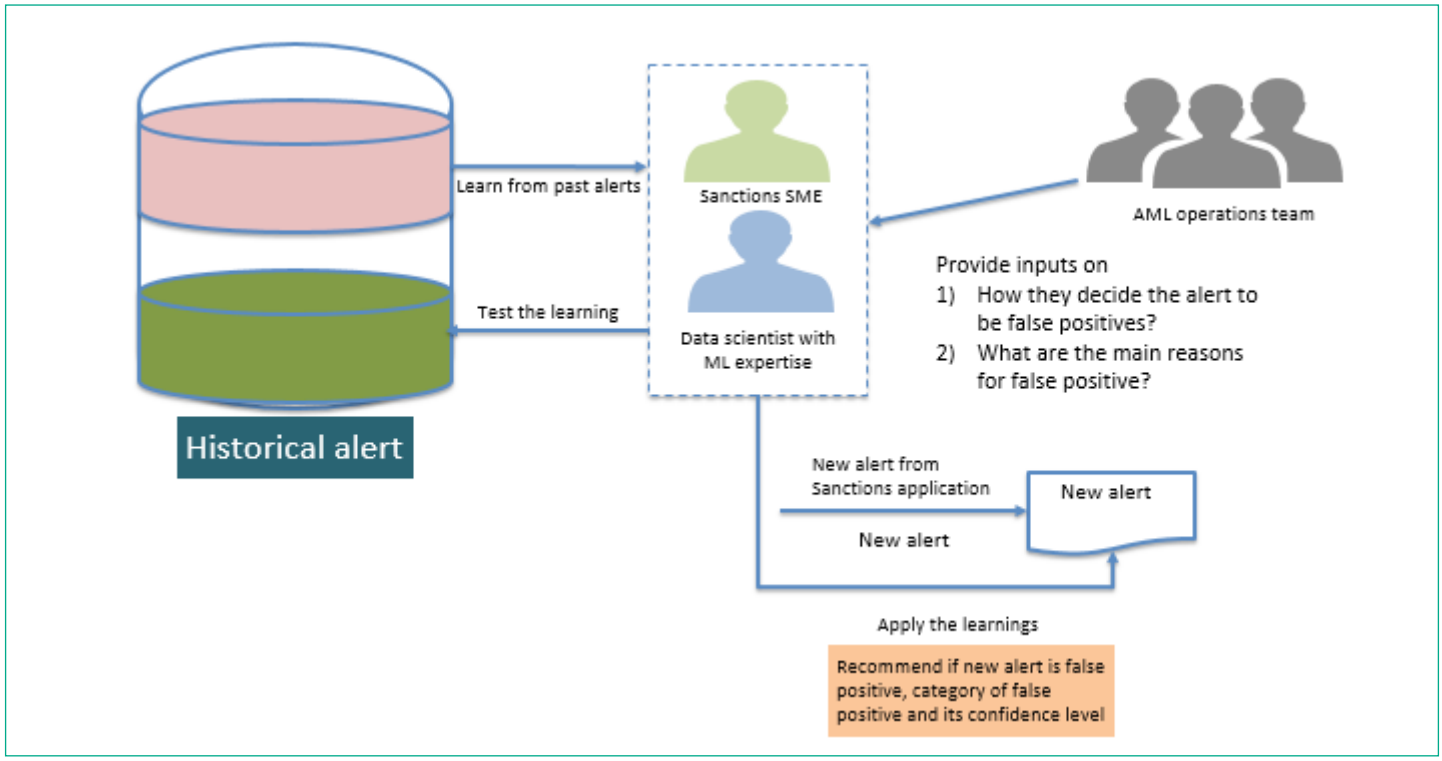


Exhibit 6: *Implementation approach*

Case Study

Infosys team was involved in a proof-of-concept (PoC) for one of the large Australian banks. The bank uses a traditional vendor's product for their sanctions screening process. However, the bank has been struggling with high number of false positives.

Existing landscape:

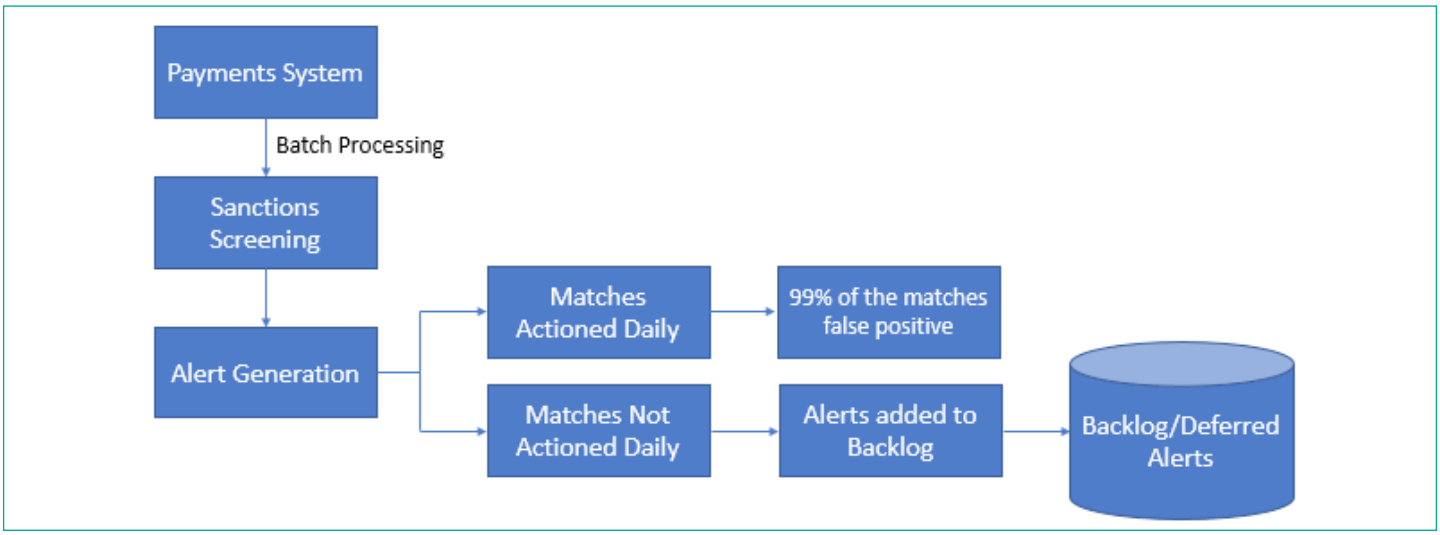


Exhibit 7: *Bank's existing sanction screening yielded very high level of false positives*

Issues:

- Due to the huge number of false positives, there are significant costs incurred by the operations team to close these alerts.
- For the matches not actioned on a daily basis, there is an increased risk of fraud since the alerts are open for a long time and there is a slower turnaround time.
- Non-standardized audit trail and comments.

Innovation done via PoC:

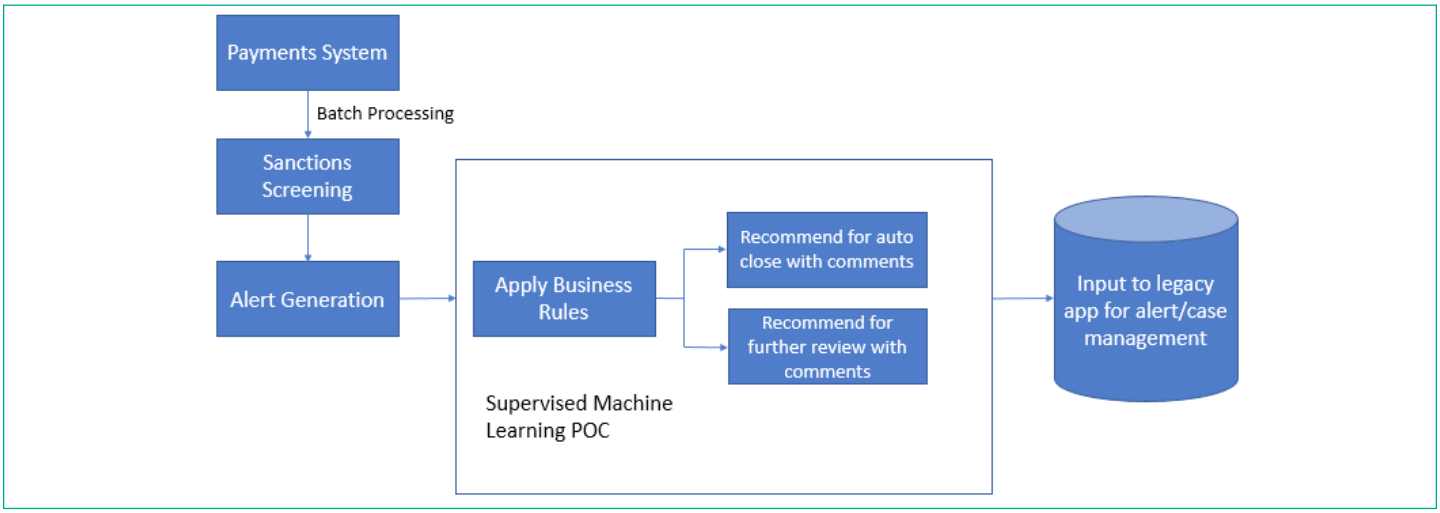


Exhibit 8: *Bank's sanction screening process post PoC*

Benefits:

- Auto closure of false positive alerts, hence, saving on effort to get them analyzed manually.
- Based on trend analysis of repetitive alerts, the main categories of false positives can be created. ML algorithm will learn and suggest auto closure of similar alerts with appropriate reasons so that operations team can take quick decisions.
- Standardized audit trail for better traceability.
- Faster and reliable decision making.
- Estimated cost savings of 60-70%.

Due to the above benefits, the bank was extremely satisfied with the PoC and has planned to go ahead with the implementation. In the current phase, the deferred alerts are being remediated through the AI/ML Model.

Conclusion

AI/ML based solution can be leveraged by FIs to effectively manage their various AML transaction filtering processes. Even where an FI is reluctant to totally replace their legacy AML transaction filtering systems with new-age AI/ML based solution, they could consider leveraging AI/ML capabilities as plug-in. The benefits from leveraging this are too substantial to ignore.

Anjani Kumar

Principal Consultant, Risk and Compliance Practice, Financial Services Domain Consulting Group, Infosys Limited

Anjani has over 20 years of comprehensive IT, domain and process consulting experience. Currently, he manages several strategic initiatives including the Competency Development Program and Thought Leadership showcasing efforts. Over the years, he has provided consulting services and managed many large and critical IT engagements for numerous key clients. He was also recognized as the lead process auditor for the IT division of a major global bank. He has extensive techno-functional skills and an in-depth understanding of quality and process models – CMMI, Six Sigma, ITIL, etc.

He can be contacted at anjani_kumar@infosys.com

Preeti Vishwanath

Lead Consultant, Risk and Compliance Practice, Financial Services Domain Consulting Group, Infosys Limited

Preeti has been with Infosys for over 16 years. During her tenure she has worked in multiple projects with US based investment banks and major banks in Australia. She has worked in both delivery and practice streams. She is currently working with a major Australian bank for the past 5 years in AML/KYC space. She has worked on various domains like Credit Risk, Prime Brokerage, Margin Risk, Collateral Management, Anti-Money Laundering, KYC, FATCA and CRS. She actively contributes to RFPs, conducts domain trainings and participates in interviews and hiring.

She can be contacted at Preeti_Vishwanath@infosys.com

References

1. https://www.baselgovernance.org/sites/default/files/2019-02/basel_aml_index_10_09_2018.pdf
2. https://rm.coe.int/moneyval-annual-report-2017-eng/16808af3c2
3. https://www.ft.com/content/71993cc2-20a9-11e9-b126-46fc3ad87c65
4. https://www.enigma.com/blog/trend-watching-across-fincens-suspicious-activity-data
5. https://www.globenewswire.com/news-release/2019/04/26/1810427/0/en/Global-Anti-Money-Laundering-Software-Market-Forecast-to-2024-Projected-to-Reach-1-9-Billion.html
6. https://www.alliedmarketresearch.com/anti-money-laundering-software-consumption-market
7. https://www.pwc.com/us/en/anti-money-laundering/publications/assets/aml-monitoring-system-risks.pdf
8. https://www.napier.ai/post/6-ways-to-reduce-false-positives-in-sanction-screening

For more information, contact askus@infosys.com

Infosys®
Navigate your next