FALSE POSITIVES REDUCTION

A case study for the reduction of false positives in the customers and suppliers screening process for a Global Blue-chip Company

NAPIER
INTELLIGENT COMPLIANCE

Introduction

BACKGROUND

A global multinational conglomerate was undergoing a major multi-year compliance transformation programme which included enhancing the sanction screening processes for suppliers and customers.

As part of the programme they had to screen millions of suppliers and customers against circa 250,000 entities in multiple sanction lists across multiple jurisdictions.

PROBLEM STATEMENT

A legacy screening solution was used to perform the screening process to identify sanctioned entities, generating circa 5% of false positives matches.

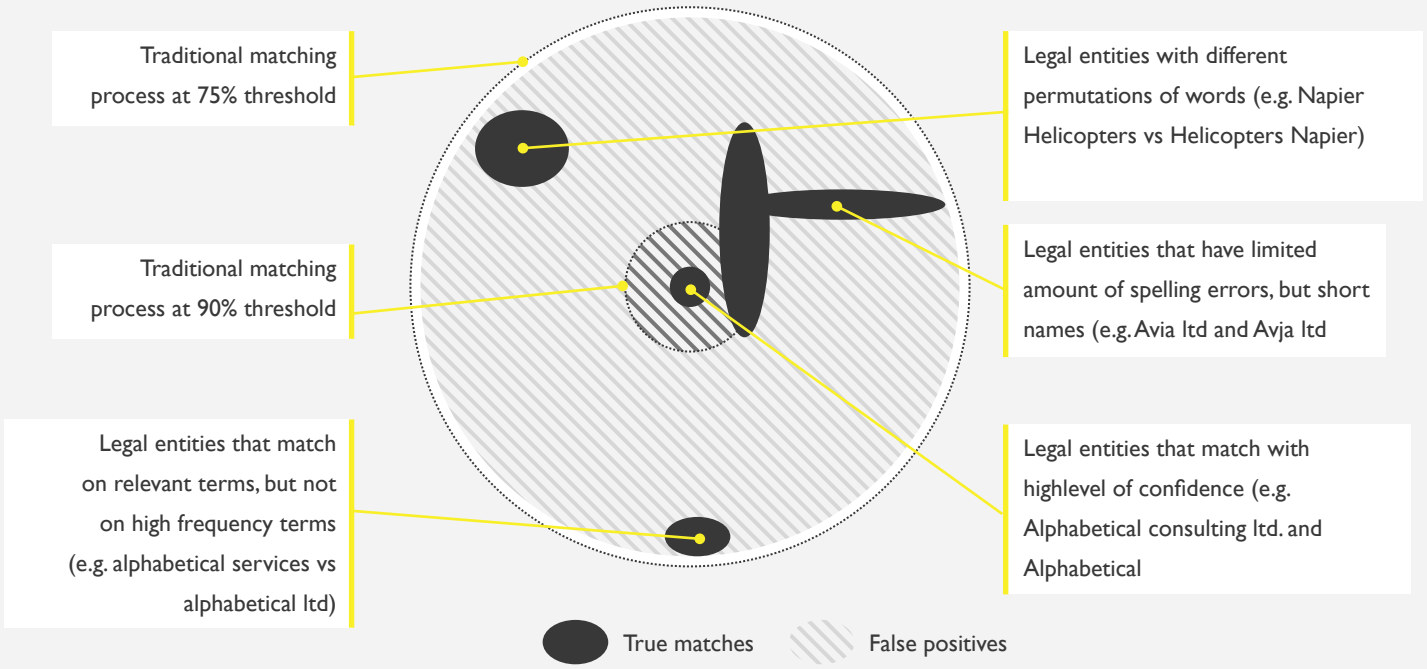
This accounted for more than 100k entities that had to be manually investigated and cleared, with multiple hits per entities.

The manual process leads to a significant screening cost (including resourcing and staffing a screening team dedicated to screen the entities for sanction risk)

OBJECTIVES

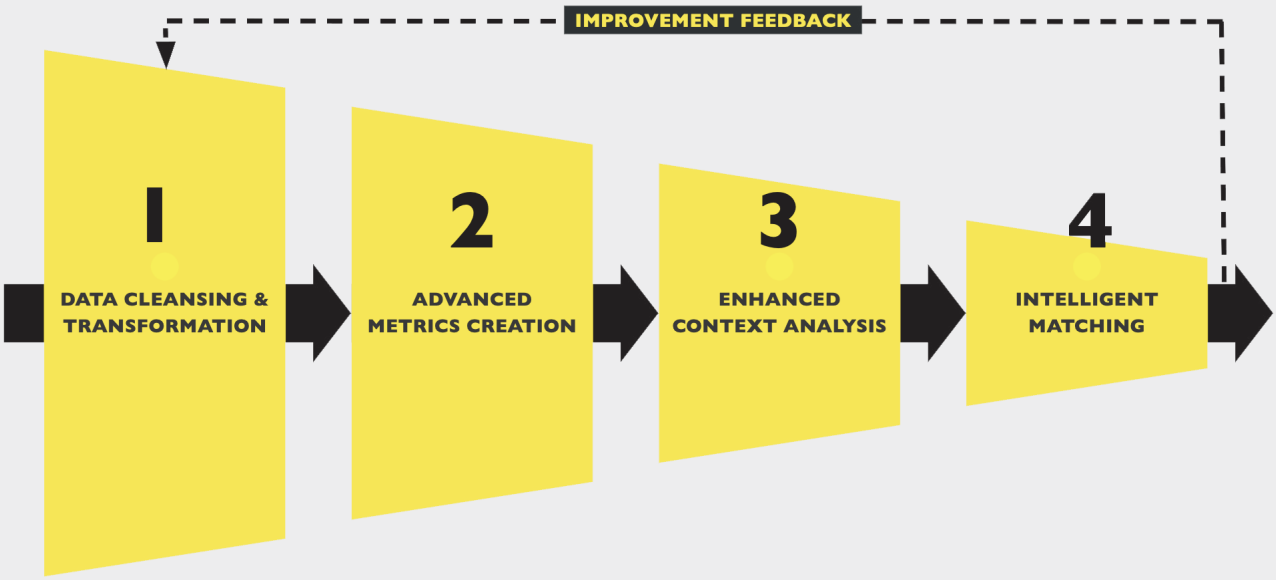
Reduce hits to be reviewed based on clients policies eliminating more than 90% of false positives thus dramatically reducing operating costs for screening hits.

Root cause for false positives



- Legacy screening systems have limitations due to the process used to screen entities (customers and vendors) and due to the algorithms used
- They require a long time to run (e.g. more than 3 days to screen 1m entities)
- Black box algorithms require thresholds to be set carefully with thorough testing to avoid false negatives
- Difficult to configure client specific policies; legacy applications are often set-up using a trial-error approach (i.e. set thresholds and check if known matches are identified)
- Lack of ability to enhance matching results by leveraging additional data

Napier False Positives Reduction



1. False positives resulting from legacy systems screening process are ingested; the data is then transformed and cleansed with initialfilters applied based on input type (e.g. domiciled country)

2. Metrics are defined based on type of data. Initial thresholds are set based on current policies and risk appetite; features are created to describe statistics of input data

3. Input data is analysed in detail to weight terms based on context; additional features are added to provide further detail of the statistics of the input data

4. Multiple techniques including machine learning are applied to the input data leveraging initial thresholds and labels; most relevant matches are presented for review

Summary outcomes

90%

REDUCTION OF FALSE POSITIVES COMPARED TO ORIGINAL PROCESS

Napier solution provides natural language explanation for the reasons why matches have been discounted, including link to policies or reference data used to automate the discount decision

£Ms

SIGNIFICANT EXPECTED COST SAVINGS FOLLOWING THE SOLUTION ROLL-OUT

The reduction of hits allowed the clients to save a significant amount of costs and reduce operating risks due to manual reviews of hits. Resources could then focus on high risk alerts as opposed to manually discount hits

0.6%

PERCENTAGE OF HIGH RISK HITS TO BE INVESTIGATED BY COMPLIANCE OFFICERS

Compared to more than 5% generated by the legacy system. Using Napier, the client was able to prioritise the work of their compliance team on the high risk individuals and entities

Beyond the use case

1. Enhance matching efficiency and effectiveness by using additional data from multiple systems (e.g. implement advanced matching on addresses, individual gender, background, etc.)

2. Create additional rules to match specific scenarios (e.g. match on nicknames, name variations, use phonetic variances for specific languages, etc.)

3. Add additional features to the data for example by automatically deriving gender of individuals from their given names using machine learning techniques

4. Use multiple machine learning approaches in parallel to improve AI generated scoring, implement champion vs challenger models, use advanced face recognition technology to improve matches' accuracy

5. Add suggestions for matches sent for review to improve prioritisation (e.g. score each match using machine learning to de-prioritise low probability matches, etc.)

6. Implement workflow to continuously improve the process using supervised anomaly detection and improve classification accuracy byleveraging manual feedback and manual data labeling

CONTACTS

NICK PORTALSKI - HEAD OF PRODUCT

Nick has extensive leadership experience in designing and delivering enterprise products using multiple technologies. Having worked in successful FinTech start-ups and enjoyed global responsibilities with IBM, his expertise lies in taking concepts from embryonic vision through to advanced end products.

E: nick.portalski@napier.ai

LUCA PRIMERANO - CHIEF AI OFFICER

Luca focuses on running artificial intelligence programmes partnering with universities and clients. His work is at the heart of the Fortytwo Data platform helping clients extracting insight from hidden data and automating decision making processes. He has extensive experience in decision automation, and digital transformation that he gained at Goldman Sachs, Deutsche Bank and Deloitte.

E: luca.primerano@napier.ai