AIFS0010

Written evidence submitted by Ms Ana Isabel Canhoto, Professor of Digital Business, University of Sussex

The interest in leveraging artificial intelligence (AI), particularly machine learning, in the fight against money laundering (AML) and criminal activities is growing. Proponents of using machine learning in AML emphasise its potential to identify previously undetected patterns in transaction data, all while maintaining cost-effectiveness. However, whether this potential is realised depends on the interaction between the technical and contextual factors specific to each AML programme.

In my research, I explored the AML detection at a UK-based financial institution, which I refer to as BANK. This institution is part of one of the UK's largest financial groups, with retail banking being its primary business, accounting for over 75% of the group's profits. Conducting such research is notably challenging. The development and application of algorithms within financial institutions are often kept confidential. Additionally, the sensitive nature of money laundering detection makes financial institutions highly cautious about discussing their methods for detecting money laundering or terrorism financing. While focusing on a single organisation for this report limits the diversity of observations and generalisability, it provides valuable, in-depth insights into an under-explored area. Furthermore, the use of various data sources and research methods, such as interviews, document analysis, and direct observation, provides a comprehensive understanding of the phenomenon.

My analysis reveals that, within AML profiling, two distinct phenomena require different approaches. The first involves building knowledge about money laundering schemes through descriptive profiling, while the second involves detecting money laundering activities via predictive profiling. Descriptive profiling relies on historical data and produces descriptive insights. For this type of task, supervised machine learning algorithms are most effective. On the other hand, predictive profiling works with real-time data and often draws on insights from descriptive profiling. It produces outputs that require further investigation by analysts, such as predictions of high-risk transactions. Unsupervised and reinforced machine learning algorithms are better suited for this type of analysis.

Both types of profiling, however, face a common challenge: financial service organisations' perspective is limited to the transactions they process. While some customers may use a single financial institution for all their banking, most use multiple providers, meaning each organisation only sees a part of the customer's financial activity. Since these organisations do not share customer data due to legal and strategic reasons, their datasets will always be incomplete. Consequently, they may fail to recognise the significance of certain transactions or misinterpret others.

Moreover, financial institutions lack direct insight into the reasons behind customers' behaviours, especially in the case of online transactions, which have become the norm for a large portion of the population. Without the ability to question customers directly about their actions, organisations are left to make inferences, which can be influenced by cognitive biases and limitations.

Legal requirements also demand that decision-making processes be explainable, ensuring that no customer is unfairly discriminated against. This is particularly challenging when AI systems are autonomous or when self-reinforcing feedback loops are present, as they are more prone to biases. Namely, unsupervised learning tends to offer lower explicability, as its outputs are harder for humans to interpret, potentially putting organisations at risk of non-compliance with regulations. Reinforced learning is most likely to generate feedback loops, especially if its rules were initially developed from unsupervised learning.

Beyond these general challenges, there are specific ones related to the two types of profiling. For descriptive profiling, one key issue is obtaining high-quality and relevant training data in a timely manner. This challenge is not unique to AML, but it is also seen in other fields utilising inductive modelling approaches, such as the case of machine learning being used for diagnosing SARS-CoV-2 infections. Initially, there was insufficient high-quality data for training, which hindered the

development of useful models. Another challenge is the evolving nature of criminal behaviour. Criminals continuously innovate their money-laundering techniques, using mobile payments or cryptocurrencies, which means that training datasets can quickly become outdated and lose their applicability.

For predictive profiling, challenges stem from the assumptions underlying the models, which are difficult to test before they are fully scaled. One key assumption is that most customers are not involved in money laundering, but this may not hold for every financial institution or type of crime. Another assumption is that money launderers' transaction patterns are distinct from those of other customers. Given the dynamic nature of money laundering, this is not always true. The difficulty in testing these assumptions means that validating the models is nearly impossible. In predictive modelling, analysts must also make deductions about how criminals might attempt to use the financial system to launder money, relying on their own reasoning. This process is susceptible to cognitive biases and stereotyping. These difficulties are compounded by the fact that predictive modelling in AML focuses on identifying unusual patterns within a client base rather than actual criminal behaviour. Treating flagged accounts as definitively linked to money laundering would lead to significant disruptions and customer complaints. As such, manual verification of algorithmic outputs is necessary, which introduces both delays and additional costs.

Finally, the cost of AI and machine learning technologies must not be underestimated. Smaller financial institutions may not have the budget for advanced AI solutions, while larger organisations operating in multiple jurisdictions may require standardised solutions that are difficult to find. Additionally, organisations must weigh trade-offs between processing speed, confidence in results, and learning curves. Adopting AI solutions often necessitates further investments in areas like upgrading legacy systems, recruiting specialised staff, and managing customer dissatisfaction. Therefore, assessing the cost-effectiveness of AI solutions is complex, with both direct and indirect costs to consider.

In conclusion, the short- to medium-term potential of machine learning for AML in individual financial institutions has likely been overstated in the literature. To fully harness its potential, standardised approaches to transaction monitoring must be agreed upon.

For more detailed insights, please refer to my full paper: Canhoto, A. I. (2021). Leveraging machine learning in the global fight against money laundering and terrorism financing: an affordances perspective. Journal of Business Research, 131, 441-452. DOI: https://doi.org/10.1016/j.jbusres.2020.10.012.

March 2025