Model Risk Management in the Machine Learning Era

Executive Summary

Over the past two years, TruEra has engaged with over 30 leading financial services firms world-wide, and several regulators (**UK**, **Singapore** and **USA**), in the industry's efforts to adapt Model Risk Management (MRM) practices for the responsible adoption of Artificial Intelligence (AI) and Machine Learning (ML). This white paper summarizes the key findings from these interactions.

The adoption of Artificial Intelligence and Machine Learning in Financial Services continues at pace. For example, 72% of the firms that participated in a recent **Bank of England survey** reported using or developing ML applications, with the number of such applications expected to grow by 3.5X in the next three years.

Greater adoption of ML has heightened expectations from MRM in financial services firms. MRM teams are expected to deal with specific technical concerns around ML models such as explainability, unfair bias and overfitting, and handle greater regulatory scrutiny and uncertainty. The number of models has gone up manifold, use cases are much more diverse, and models originate from many more sources, both inside and outside the firms. There is an expectation of faster turnaround and greater throughput in model reviews and approvals. These challenges are compounded by constraints in the quantity and depth of ML-focused MRM talent.

Sophisticated adopters of ML have responded to these expectations through a multi-pronged strategy. They have enhanced their MRM methodologies around explainability, fairness, and various aspects of conceptual soundness (for example: sensitivity to small changes in input data). Many have attempted to standardize and automate several aspects of model testing and reporting, as a way to complement additional, ML-focused resourcing in MRM teams. Some have introduced more frequent, automated model monitoring to provide triggers for investigation and fall-back mechanisms.

Several firms have revisited their model materiality frameworks to enable greater discrimination in the level of attention provided to each model. In a small number of instances, firms have decided to limit use of ML models in areas where regulatory or reputational risks might overshadow potential benefits.

Technology has played a key role in enabling these responses. Specifically, **TruEra's AI Quality platform has been used by clients to transform their MRM practices in four ways:**


Providing comprehensive, interactive tooling

for model developers, validators and auditors to explain and understand ML models, test them for potential weaknesses, conduct root cause analysis for identified issues and/or debug the models
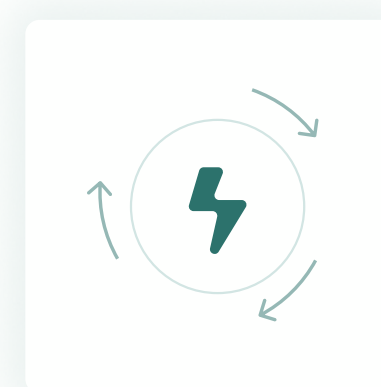

Creating a suite of standardized performance, stability, robustness and fairness tests

that automates several aspects of ML model validation, and makes self-assessment of low-materiality models by first line teams a viable option
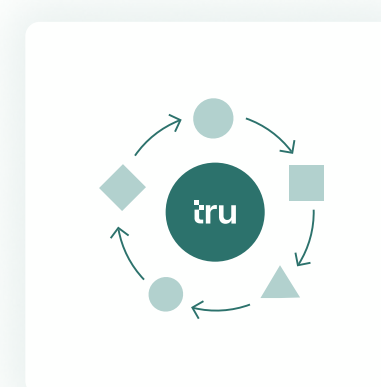
Auto-generating 30-70% of standard MRM documentation templates for ML models,

directly improving the efficiency (and quality) of the governance process
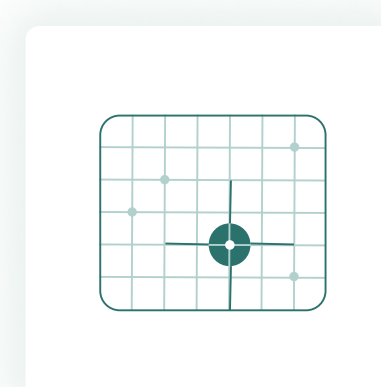
Dramatically improving their ability to monitor changes in model effectiveness

and quickly debug live ML models to trigger corrective actions

To read the full whitepaper, go to **Model Risk Management in the Machine Learning Era**.