

**Abstract:**

The use of advanced (complex non-linear) models, such as deep neural networks and boosted decision trees, has become increasingly prevalent over the last five years or so. Data scientists are routinely creating models that significantly outperform simpler traditional linear approaches within out of time samples available at the point of development. This is not a surprise as advanced models can often identify patterns and relationships that may be missed by linear models, leading to better predictions.

These models are, by definition, incredibly complex, often incorporating hundreds, if not thousands, of weights or parameters, which can make it difficult to understand how the model arrived at a particular prediction. Additionally, advanced models may be more susceptible to 'over-fitting' as evidenced by the typical gap in performance between training and holdout samples.

In contrast, traditional linear models are often simpler and easier to interpret, making them more transparent and easier to validate over the long term.

**This leads to the question: do advanced models provide a sustainable uplift over linear approaches or does performance degrade more rapidly and necessitate more frequent re-optimisation/redevelopment?**

While frequent redevelopment is less of a concern in non-regulated industries, it can be time and resource intensive in regulated industries due to the higher level of oversight and scrutiny they receive from regulatory bodies.

**Using longer run performance data, this analysis explores the subject and examines whether population instability disproportionately impacts advanced models.**