Suncor: Performance and explainability aren't always a trade-off

Summary

Deep learning often performs well at predicting processes that are nonlinear (i.e., small changes can have an outsize impact) and highly coupled (i.e., there is a lot of dependency between factors). Often there is sufficient performance improvement from adopting deep learning that firms are willing to have their models be less explainable (i.e., black box). But, as Suncor's experience shows, when the stakes are high enough, explainability is at a premium.

Opportunity: Better prediction of issues, to manage output quality

Suncor Energy specializes in the production of synthetic crude from oil sands. For Diesel in particular, this involves removing sulfur and nitrogen by hydrotreating, mixing the straight-run diesel with hydrogen (and a solid metal catalyst e.g., cobalt) at high temperature and pressure. The process is complex, with several variables (e.g., pressure, temperature, flow rates) that interact to affect the resulting quality of the output. These factors must be tightly monitored and controlled to minimize "upsets", where product quality deviates outside acceptable rates and cannot be sold. With Diesel production averaging 43000 barrels per day, there is a strong commercial incentive to avoid upsets as much as possible.

Challenge: High-impact decisions require high explainability

Ultimate accountability for the output quality rest with *Sitewide Leads*, who oversee the production sites and make key operational decisions that influence output quality. Understanding which adjustments to make and their resulting effects can often come down to individual experience and judgement. To be defensible, these high-impact decisions need to be articulated with a clear rationale, which means that any analytical technique for decision support must be transparent and well understood.

Solution: PCA had greater explainability, and comparable performance

Suncor set out to improve their prediction capability, by developing an "upset flagging" model whereby emerging suboptimal conditions could be identified with enough time for corrective action to be taken. They identified 11 factors associated with product quality that needed to be assessed in real-time and built models that incorporated 30 different measurements from sensor data. Initially, the data science team explored a wide range of sophisticated, but hard-to-interpret machine learning techniques, including Neural Networks, LSTMs, Random Forests, Gradient Boosting, and Decision Trees, with XGBoost (an ensemble technique combining Decision Trees with Gradient Boosting) performing the best.

However, when they compared the performance to simpler (e.g., traditional statistical) techniques, they observed performance that was much better than expected. For example, Principal Component Analysis (PCA) performed at only a 10% performance deficit relative to XGBoost, while being much easier to interpret.

Outcomes: Early warning with transparent, defensible rationale

After thoroughly testing both approaches with their Sitewide Lead stakeholders, Suncor decided that the greater explainability more than outweighed the performance trade off in this case. Together with the predictions, using PCA enabled a read-out of the associated weightings behind each factor i.e., a ranking of which factors were most important in driving the prediction. The resulting system was able to predict upset events up to an hour in advance, with an accuracy of 80%, every 5 minutes.