CAUSA ENS







Causal Al

Why Causal AI?

Current machine learning approaches have produced remarkable achievements for certain types of data such as images and language - where the training data is plentiful, and the underlying 'data generating process' does not vary over time. For example, it takes many years for a language to evolve and a cat always looks like a cat.

However, there has been surprisingly little progress for data types such as time-series, which are ubiquitous in finance and business. The problem is that current machine learning approaches 'overfit' the data - they attempt to learn the past perfectly, instead of uncovering the 'real', or causal, relationships that will continue to hold over time.

"Current machine learning techniques are unable to understand cause and effect"

At present, predictive models for time-series are mostly curve fitting exercises. As a consequence, models are driven by parameters that happened to correlate in the past but may not be capable of predicting the future.

At causaLens we believe that a new theory of how to build intelligent machines is required. Machines need to be capable of understanding "cause and effect" in order to advance machine learning and to make reliable predictions of revenues, stock prices or real estate yields.

Case Study

Discovering Predictive Signals

To demonstrate how Causal AI technology works in practice we present the following case study, which illustrates how our technology autonomously discovered value in shipping data. The value was discovered thanks to the technology's ability to differentiate between spurious correlations and causality in the dataset.

Real-time Shipping Data

MariTrace

MariTrace systematically monitor the movements of over 50,000 commercial vessels in order to understand the flow of commodities around the world in real-time.

Using weekly liquid natural gas (LNG) shipping exports (a fraction of total natural gas exports), our technology autonomously found a signal to predict the price of UK natural gas.

Results with Current Machine Learning

Standard machine learning techniques, including traditional AutoML, would have discovered that the price of natural gas is correlated in a significant way to the following variables:

- Qatar LNG exports
- Trinidad & Tobago LNG exports

The figure below shows the correlation between LNG exports for different countries and UK natural gas price. The green bars represent those with correlations which are statistically significant.



 \boxtimes



These statistically significant correlations would then be used to drive forecasting models. However, using correlations that are not causal would result in 'overfitting' and reduce the performance of the models in production, as shown in the section below.

Results with Causal AI

Current approaches often result both in false positives, identifying drivers that are not predictive; and in false negatives, failing to identify predictive drivers. In contrast, our Causal AI technology is specifically designed to uncover the true causal relationships in data.

Applying this technology we found that Qatar LNG shipping exports are a stable predictive signal for UK natural gas futures, while also identifying that the correlation between Trinidad & Tobago shipping exports is spurious.

To show that the causal relationships identified lead to improved models in the real world we built thousands of models. Half of the models were discovered by Causal AI and the other half were built using standard machine learning techniques. Both sets of models were then given data they hadn't previously seen to make new predictions of the price of natural gas and assess their performance.

Models built with Causal AI had an average model score 42% higher, including 13% higher directional accuracy, than those built with current standard methods. This study demonstrates how using Causal AI to identify true causal drivers can result in better models that will maintain their performance in production.

Summary of Findings

- Qatar LNG shipping exports were found to be a stable predictive signal for UK natural gas futures price.
- Current machine learning technologies based on correlation analysis are insufficient to discover value in data. For example, Trinidad & Tobago shipping exports would have been incorrectly selected leading to suboptimal models.
- Models produced with Causal AI provide superior results, reducing overfitting.

Unlocking Potential with Causal AI

Autonomously identifying the causal relationships that affect the price of natural gas is a simple example of using causal AI for time-series. More complex problems of this nature are widespread in many industries, such as finance, logistics, transport, energy, retail and healthcare. The application of Causal AI is already unlocking previously undiscovered value and can be used to optimise any business at scale.

"Causality holds the key to unlocking the true potential of AI"



causaLens model score

About Us

causaLens leads causal AI research: a way for machines to understand cause and effect. We believe that causality is a major step towards developing true AI and fulfilling our vision of optimising the global economy. The causaLens platform is becoming the industry standard in financial services and is increasingly being used in other industries. causaLens is run by top scientists and engineers, 70% holding a PhD in a quantitative field. Contact us on info@causaLens.com or follow us on LinkedIn and Twitter.