

Swissquote Conference 2019 on Artificial Intelligence in Finance

Jan Witte (UCL)

Title: Synthetic Financial Data: An Application To Regulatory Compliance For Broker-Dealers

Abstract: The hype of “big data” has not escaped the investment management industry, although the reality is that price data from U.S. financial markets are not really big data; price data is small data. The fact that sellers and advisors in financial markets use small data to generate and test investment strategies creates two major problems. First, the economic mechanisms that generate prices (and, therefore, returns) may change through time, so that historical data from an earlier time may tell us little or nothing about future prices and returns. Second, even if data-generating-mechanisms are somewhat stable through time, inferences about the profitability of investment strategies may be sensitive to a handful of outliers in the data that get picked up again and again in different strategies mined from the same small data set. We present an answer to the financial small data problem: using machine-learning (ML) methods to generate “synthetic” financial data. The essential part of our approach to developing synthetic data is the use of ML methods to generate data that might have been generated by financial markets but was not. Synthetic price and return data have numerous uses, including testing new investment strategies and helping investors plan for retirement and other personal investment goals with more realistic future return scenarios. We focus on a particularly important use of synthetic data: meeting legal and regulatory requirements such as best interest and fiduciary requirements.