

Predictive Analytics: Theory, Applications and Algorithms

How can policy makers and businesses adapt and confront the rapidly-evolving world through predictive analytics? The International Institute of Forecasters (IIF), in collaboration with Anglia Ruskin University, Siemens AG Munich, the German Operations Research Society, and Hughes Hall successfully organized an international workshop on 23-24 May 2019, titled "Predictive Analytics: Theory, Applications and Algorithms." The two-day workshop provided a unique forum for practitioners and researchers to discuss and evaluate the methods through cutting-edge research. The agenda embraced the need for benchmarks and the use of test-training-validation sets in predictive analytics as well as the design and application of novel algorithms using machine learning and AI to prediction and causal impact analysis.

'Predictive analytics' is used to include a variety of different techniques, such as econometric modelling, Machine and Deep Learning, and Artificial Intelligence (ML, DL and AI), to improve efficiency and reduce risks at macro and micro level through analyzing, predicting and forecasting operational data, optimizing supply chain management and marketing campaigns, detecting fraud and hidden sources of market power and bottlenecks, etc.

Opening the workshop with Emanuele Giovannetti, Mohsen Hamoudia from Orange Labs and Lyon Business School presented briefly the various activities of the IIF, and the IJF and Foresight journals.

Utilizing new mobile payment data, Robert Fildes from Lancaster University Centre for Marketing Analytics and Forecasting proposed a third party mobile-payment platform centered customer flow forecasting solution based on an extension of the newly developed Gradient Boosting Regression Tree (GBRT) method the results clearly demonstrated the potential strength of ML methods. Evengelios Spiliotis from the National Technical University of Athens challenged the use of ML methods in business forecasting, suggesting a hybrid method that mixed Neural Networks (NN) with exponential smoothing formulas and urged us to identify best practices for applying appropriate data-preprocessing, cross-learning in a smart combination of with machine learning to achieve accurate time series forecasting, as the M4 Competition has demonstrated. Larry Vanston from Technology Futures Inc further addressed the speed of the adoption of AI in predictive analytics and forecasting, focusing on identifying the drivers of adoption and the various constraints these new methods face: a forecast of AI's use in forecasting

Rohan Kariyawasam from Anglia Ruskin University posed two important questions in his talk on the elusive test of dominance in the digital economy, 1. Do we need to redefine significant market power in the digital economy? 2. Could the EU regulatory framework offer loopholes that undermine the regulation's effectiveness? He developed a layered policy model to more accurately define the internet market, providing regulators with a framework to assess Next Generation Networks more accurately. Whilst Claudio Piga from Keele University showed how the main features of a pricing algorithm for airline tickets were shaped by organizational frictions, which impede the continuous price adjustments recommended by revenue pricing models.

Causal trees and causal forests models were proposed in new developments in predictive analytics by Melvyn Weeks, Eoghan O'Neill and Jinhua Wang from the University of Cambridge. Their aim was to estimate individual estimates of heterogeneous treatment effects. Melvyn Weeks applied causal forest algorithm on Time-of-Use pricing schemes on 4,225 household electricity consumption, showing how algorithm can be used to target consumers in electricity market. Eoghan O'Neill extended an alternative Bayesian Additive Regression Trees fitting algorithm by adopting Bayesian Causal Forests using Bayesian Model Averaging (BCF-BMA) to examine the treatment effect. An alternative approach was presented by Jinhua Wang who criticized tree models and proposed honest estimation through further exploration of the heterogeneity of treatment effects with Causal Trees.

Whilst tree models generated heated debates on the effectiveness of models and benchmarks, Ralph Grothman from Siemens discussed how Siemens applied deep feedforward neural networks to forecast customer demand through data analysis and AI on predictive analysis and diagnostic analysis for planning and production, which has been successfully reduced 20 percent errors in forecasting more than 1,000 products. Emanuele Giovannetti from Anglia Ruskin University adopted Artificial Neural Networks to capture the emergence of bottlenecks in the upstream internet connectivity networks. With a focus on algorithmic prediction of interconnection structures and its impact on market power, his research showed its practical implications on the internet market, indicating competition authorities should build prioritization policies based on AI. William Davies explored the impact of social capital on crowd-funded lending projects on the KIVA platform, a not-for-profit organization focused on very small business opportunities. Using a backwards propagating neural network model, the analysis showed how the network eigenvector centrality, that captures network and project relationships, affects crowdfunding through social capital. The use of ML proved more informative than standard regression approaches.

Martin Bellamy, Chief information officer from Financial Conduct Authority, UK, introduced how FCA applied Machine Learning and AI to improve forecasting and projection to detect market abuse and irregular market behavior, indicating advanced predictive analytics as their prime tool to ensure effective market operations. Adopting artificial neural network on high frequency financial trading data, Arze Karam from Durham University explored the same theme. She provided evidence that algorithmic traders adopted specific types of strategies, such as undercutting and quote stuffing, to increase quote volatility and execution costs, significantly affecting the stability of financial markets and increasing the frequency of market crashes. Meanwhile, Aurelio Bariviera from Universitat Rovira i Virgili examined cryptocurrency investment using dynamic linguistic information aggregation approaches, showing the capabilities of the dynamic decision models as a tool to support decision making in financial investment.

To conclude, the workshop's success was based on the variety of perspectives, both theoretical and in application, taken by the participants. The final discussion asked the question, 'What can ML methods offer beyond established approaches' to the range of problems areas, both cross-sectional and time series. Crucial issues remain identifying appropriate benchmarks to compare with ML results and the split between training, validation and test data. Only then can we be certain that gains are being achieved.