INSEAD
The Business School for the World®

KNOWLEDGE

Decision Makers Should Rely on Hybrid Forecasting Models

Machine learning may be hyped as the way of the future, but as a forecasting method, it works best when combined with standard algorithms.

Just about every business depends on accurate forecasting. A classic example is a manager forecasting the amount of goods to produce or the level of inventories to keep. In the case of Uber, the ride-hailing firm requires sophisticated models to predict ride supply-and-demand, as well as how much personnel is needed for its customer service and support systems. Even hardware requirements can be predicted: Under-provisioning may lead to outages, but over-provisioning can be very costly. Forecasting can help find the sweet spot.

Since forecasting is at the heart of its operations, Uber has invested heavily in building solid related expertise. It employs data scientists who keep up with cutting-edge techniques in machine learning, probabilistic programming and other methods to ensure the accuracy of its forecasting algorithms.

Recently, one of Uber's leading data scientists, Slawek Smyl, beat 48 individuals and teams from countries around the world to win the M4 Competition, the fourth and latest edition of the Makridakis Forecasting Competitions (M Competitions). The purpose of these competitions is to guide organisations on how to improve the accuracy of their predictions and assess future uncertainty as realistically as possible. Along with practitioners and academics, major firms such as Oracle and Wells Fargo participated in the event.

Having started at INSEAD more than 40 years ago and held roughly every decade since, the M Competitions compare the accuracy of various time series forecasting methods. In the field of decision sciences, *time series forecasting* essentially means to predict future values based on past ones observed at regular time intervals. Tide heights or the daily closing value of a stock exchange are examples of such observable data. The latest competition covered six application domains (macro, micro, demographic, industry, financial and others) and six time frequencies (from hourly to yearly).

In the first three M Competitions, simple forecasting methods did well, while in the M4, sophisticated ones provided a significant edge for improving forecasting accuracy. The M4 Competition used a very large data set – 100,000 time series – and the results confirmed that pure machine learning and neural network methods performed worse than standard statistical methods, and worse still than various combinations of such methods.

The major findings of the M4 Competition

As a whole, the competition's results show that both statistical and machine learning methods are of

limited value when taken in isolation. When it comes to improving forecasting accuracy and making forecasting more valuable, hybrid approaches and combination methods are the way forward.

Described in greater detail in **"The M4 Competition: Results, findings, conclusion and way forward"**, published in the *International Journal of Forecasting*, the five major findings of the M4 Competition are as follows:

1. The combination of methods was the king of the M4. Of the 17 most accurate methods, 12 were combinations of mostly statistical approaches.
2. The biggest surprise was a hybrid approach combining statistical and machine learning features, which was nearly 10 percent more accurate than the combination benchmark. Submitted by Smyl, this method produced both the most accurate forecasts and the most precise prediction intervals.
3. The second most accurate method was a combination of seven statistical methods and a machine learning one. The averaging weights were calculated by a machine learning algorithm trained to minimise forecasting errors through holdout tests. This method was submitted jointly by Spain's University of A Coruña and Australia's Monash University.
4. The first and second most accurate methods managed to correctly specify the 95 percent prediction intervals, an amazing success in itself. These are the first methods we are aware of that have done so. Typically, forecasting methods tend to considerably underestimate uncertainty.
5. The six pure machine learning methods entered in the competition all performed poorly.

The **M4 Conference** will be held on 10-11 December 2018. Speakers from major tech companies (Google, Microsoft, Amazon, Uber and SAS) and top academics will meet in New York City to elaborate on the findings of this year's competition. The developers of the three most accurate methods will explain how business and other organisations could apply them. Keynote speakers include Nassim Nicholas Taleb, author of *Black Swan* and *Skin in the Game*, who will talk about uncertainty in forecasting and how he thinks that tail risks are worse today than they were in 2007, just before the Great Recession.

Spyros Makridakis *is an INSEAD Emeritus Professor of Decision Sciences and Professor at the University of Nicosia.*

Slawek Smyl *is a Data Scientist at Uber Technologies.*