Mind the Gap?

Hype-Cycle versus Business Reality in Data Science (a forecasting perspective)

Sven F. Crone

16/09/2016
Agenda

Mind the Gap?

- The Hype?
  - Gartner’s Hype Cycle over time

- The Gap?
  - Is there a Gap?
  - If so, how large is it?

- How to close the Gap?
  - More research?
  - More applications?

Emerging Technologies Hype Cycle

New Technologies make bold promises
→ how to discern the hype from what’s commercially viable?

Should you make an early move?
Take a moderate approach?
Wait for further maturation?

Mind the Gap in Data Science

©2016 Sven F. Crone | All rights reserved.
We are on the plateau of productivity!

In 2015 Gartner dropped Big Data, Analytics, and Data Science!

Are we all on the plateau of productivity?
Hype Cycle for Advanced Analytics

Figure 1. Hype Cycle for Advanced Analytics and Data Science, 2016

looks like we are back in the hype! (and have 5 more years)


The average 2016 respondent used 8.1 algorithms, a big increase vs a similar poll in 2011.

unclear which TS algorithms are used?
Brown (1956) and Box & Jenkins (1970) predicted that the trend would continue "for at least ten years." Moore (1965) predicted that the number of components in integrated circuits (IC) has doubled every year from its invention in 1958 until 1965.


Mind the Gap?

- The Hype?
  - Gartner’s Hype Cycle over time
- The Gap?
  - Is there a Gap?
  - If so, how large is it?
- How to close the Gap?
  - More research?
  - More applications?
**Evidence from Forecasting Practice?**

**Practitioner Survey**
- Target group: demand planning & forecasting professionals (in manufacturing)
- LCF Mailing list, forecasting lists / blogs (ISF, SAS), 100s of LinkedIn Groups, 2000+ personalised LinkedIn invites

→ 540 responses (representative)
→ 200 valid surveys in manufacturing

**Robust Questionnaire Design**
- Pilot study in 2011 (to ensure validity)
- Final pre-version pre-tested with 18 FMCG forecasters
- Conducted January 2012-August 2012
- Validity for large manufacturers active in the FMCG / CPG industry

Evidence from Forecasting Practice?

1. Use of judgment introduces (cognitive) biases
2. Use of judgment limits automation!

(i.e. “Statistics” aka Data Science or “Gut-Feel”?)

What approach do you use in forecasting?

70% of industry forecasters employ some form of judgment
Evidence from Forecasting Practice?

Use of methods developed & implemented in the 1950s-1960s
Does your vehicle fleet look like this?

[Image of vintage cars]

Austin Healey 3000 (1959)

BMW Isetta (1956-1962)

Does your hardware look like this?

[Image of vintage computer equipment]

Working with IBM type 704 electronic data processing machine used for making computations for aerodynamic research.

21 March 1957, NASA
Do your employees look like this?

Then why do your algorithms look like this?

\[ L_t = \alpha \cdot Y_t \cdot S_t \]  
\[ T_t = \beta \cdot L_{t-1} + (1-\beta) \cdot T_{t-1} \]  
\[ S_t = r \cdot \frac{F_t}{L_t} + (1-r) \cdot S_{t-1} \]  
\[ F_{n+k} = (1 + r) \cdot F_k \cdot S_{n+k} \]
\[ L_t = \alpha \cdot Y_t \cdot S_t - s + 1 - \alpha \cdot L_{t-1} - 1 + T_t - 1 \]
\[ T_t = \beta \cdot Y_t \cdot L_t + L_t - 1 + 1 - \beta \cdot T_{t-1} \]
\[ S_t = \gamma \cdot Y_t + 1 + \gamma - B \cdot S_{t-1} \]

Enos in Mercury-Atlas

Juri Gararin in space

US in Pigs's bay

"It can drive a 6-D nail thru a 2 X 4 at 200 yards"
How to measure "algorithm maturity"?

\[ \text{dog has } 10 \text{ years life expectancy} = 5.6 \text{ generations} = 403 \text{ dog years} \]

IT-Hardware §253 Abs. 3 Satz 1 HGB
- Mainframes: 7 years depreciation
- Workstations: 3 years
  - 11.2 generations
  - 784 IT years
Mind the Gap in Data Science

©2016 Sven F. Crone | All rights reserved.

Mind the Gap?

- The Hype?
  - Gartner’s Hype Cycle over time
- The Gap?
  - Is there a Gap?
  - If so, how large is it?
- How to close the Gap?
  - More research?
  - More applications?

Agenda

- The Hype?
- The Gap?
- How to close the Gap?
3900 papers on Neural Networks
(only in forecasting, 117,923 in total)
327,693 in topic

377 papers on Exponential Smoothing
5,657 in topic

Upgrade to 1970s technology ... ?

ARIMA
SARIMAX
Multiple Regression
Forecast Combination (Ensembles)
Upgrade to 1980s technology ... ?

- Advanced Exponential Smoothing
- GETS Multiple Regression, SCANPRO
- Double & Triple Seasonality ETS
- Bootstrapping, zero-inflated demand

Upgrade to 1990s technology ... ?

- State space models
- Combination (Ensembles)
A Network of Neurons

Multilayer Perceptrons

- Class of statistical methods for (nonlinear auto-)regression
- Combination of simple processing units \rightarrow complex system behaviour

\[ \hat{y}_{t+h} = f(x_t, x_{t-1}, \ldots, x_{t-n}) + \varepsilon_{t+h} \]

Each MLP calculates a NARX(p)-process
Mind the Gap in Data Science

©2016 Sven F. Crone | All rights reserved.

Fit appropriate forecasting models to:

- Predict future value
- Understand effect (elasticity) of each exogenous factor \(\rightarrow\) affect forecasts

\( f(X) = +125.37\times \text{Constant} + 0.71\times \text{Lag1} + 0.07\times \text{Lag2} + 610.76\times \text{Adcode: 401} + 617.24\times \text{Discount} - 1569.46\times \text{Xmas} - 696.48\times \text{Xmas+1} + 1209.29\times \text{Xmas-2} - 1581.59\times \text{Easter-1} - 894.91\times \text{Labour} \)
Reality is highly nonlinear, especially when considering multiple variables. Neural networks (and other advanced nonlinear causal models) are a way forward.

Simulated time series of AR(3)X with weather & price variables
Models nonlinear interactive Demand-Price Elasticity

Improved Forecasting Algorithms

<table>
<thead>
<tr>
<th>Model</th>
<th>Train</th>
<th>Valid</th>
<th>Test</th>
</tr>
</thead>
<tbody>
<tr>
<td>Seasonal Linear Regression (35B)</td>
<td>38.41609</td>
<td>16.61117</td>
<td>18.20583</td>
</tr>
<tr>
<td>MLP AR, SinCos + Selection</td>
<td>17.82759</td>
<td>11.73440</td>
<td>9.13354</td>
</tr>
<tr>
<td>Improvement</td>
<td>-20.5885</td>
<td>-4.87677</td>
<td>-9.07229</td>
</tr>
<tr>
<td>Improvement in %</td>
<td>-53.59%</td>
<td>-29.36%</td>
<td>-49.83%</td>
</tr>
</tbody>
</table>
### Improved Model Selection: Human expert vs. system?

<table>
<thead>
<tr>
<th>Country</th>
<th>Prior Model</th>
<th>IF Model</th>
<th>Change % points</th>
<th>Change in %</th>
<th># items</th>
</tr>
</thead>
<tbody>
<tr>
<td>Canada</td>
<td>40.7</td>
<td>33.8</td>
<td>-6.9</td>
<td>-16.9%</td>
<td>47</td>
</tr>
<tr>
<td>Germany</td>
<td>55.4</td>
<td>51.7</td>
<td>-3.7</td>
<td>-6.8%</td>
<td>155</td>
</tr>
<tr>
<td>France</td>
<td>43.7</td>
<td>42.6</td>
<td>-1.2</td>
<td>-2.7%</td>
<td>262</td>
</tr>
<tr>
<td>Greece</td>
<td>50.9</td>
<td>49.4</td>
<td>-1.7</td>
<td>-3.3%</td>
<td>196</td>
</tr>
<tr>
<td>Italy</td>
<td>42.7</td>
<td>39.9</td>
<td>-2.8</td>
<td>-6.5%</td>
<td>175</td>
</tr>
<tr>
<td>Netherlands</td>
<td>41.0</td>
<td>38.9</td>
<td>-2.1</td>
<td>-5.1%</td>
<td>154</td>
</tr>
<tr>
<td>Poland</td>
<td>55.2</td>
<td>47.1</td>
<td>-8.1</td>
<td>-14.7%</td>
<td>78</td>
</tr>
<tr>
<td>South Africa</td>
<td>37.3</td>
<td>35.9</td>
<td>-1.4</td>
<td>-3.7%</td>
<td>36</td>
</tr>
<tr>
<td><strong>Average</strong></td>
<td></td>
<td></td>
<td><strong>-7.5%</strong></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- Automatic Model Selection able to improve slightly on Expert
- Automatic Model Selection reduces work significantly

### New Product Clustering

- All time series with up to 27 months of observations
- Hard to extract useful information
- No additional information
- Focus on the first three months (new products)

⇒ How to define if things “look” similar?
1. Outlying daily patterns
   - Functional outliers - in level & shape
   - Predictable reasons: Bank Holidays
   - Unpredictable: Natural disasters, hurricanes, Industrial actions, etc

2. Outlying Spikes within daily pattern
   - Special occurrences during normal day

Newer (& better?) algorithms exits!
Extend your Tool Box!

Artificial Neural Networks, Support Vector Regression

C4.5, k-Means, SVM, Apriori, EM, PageRank, AdaBoost, kNN, Naive Bayes, and CART

Disclaimer
Old Tools!

Ok for some tasks!

World's largest tree-house – 10 stories tall
New Tools!

Useless ... for the task!

New Tools!

Useless.
Beware!

“Buying the most expensive Hammer ...

...does not make you the best Carpenter!”
APO DP: 78% (45% ETS + 22% Av + 11% Naïve) use simple methods!

APO drives use of simpler methods than other software packages.

Study Results (all industries)
Use of Forecasting Methods

- Exp. Smoothing is the single most used method.
- Specialist software drives use of complex methods.
- Only marginal use of Linear Regression.
- Excel drives use of even simpler methods (low-cost).
- APO DP drives use of simpler methods.

Ask yourself ...
Can you afford to ignore 60 years of progress?
Artificial Intelligence meets Forecasting.

Enhanced algorithms & model selection for SAP APO

Predictive Analytics World London, the leading vendor-neutral analytics conference, is holding its fifth annual conference this October 12-13 in London, UK at etc venues, 300 Aldersgate. PAW focuses on concrete examples of deployed predictive analytics. Join PAW London to learn exactly how top practitioners deploy predictive analytics, and the business impact it delivers.

Featured Speakers:

- Dr. John Elder
  Chief and Founder
  StatSoft Research, Inc.

- Michael Biscoe
  Analytics Director
  Thames Water Utilities

- Phoebe Clarke
  Senior Data Scientist
  Channel 4

- Sue Croce
  Director
  Lancaster Research
  Centre for Consulting

- Anti Cepas
  Director, Global Product Marketing
  Visa

- Alfie Hetherington
  Global Head of Food
  Microbiological Disease
  Prevention

- Gerhard Hone
  Director Data Services
  Siemens AG

- Simon Nourse
  Senior Data Science Manager
  GSK
Take aways

• The hype is real
  – ... but it’s just a hype you can use!

• The gap is substantial.
  – Companies are slow adopters to algorithms
  – (not all) Software companies are innovators

• The potential is substantial
  – High opportunities from low-cost pilot studies
  – Try new algorithms!
    • Neural Networks
    • Support Vector Regression
    • Decision Trees
    • K-Nearest Neighbours
    • ...