LIMITATIONS OF PREDICTIVE ANALYTICS

James OGUNLEYE, PhD

Professor at Middlesex University and the Editor of the International Journal of Developments in Big Data and Analytics, UK

Middlesex University and the Editor of the International Journal of Developments in Big Data and Analytics, UK

Outline

- Context – the phenomenon of Big Data

- Limitation of Predictive Analytics

- Minimising Limitations

- Concluding Remarks

Context: the phenomenon of Big Data

- Predictive analytics has been transformed by the phenomenon of big data.

- These datasets—85% of which are non metric data or unstructured (SAS Institute, 2012)—are huge and complex in volume, velocity, variety, veracity and variability and threatens traditional computing architectures (Ogunleye, 2014).

Conceptualising Predictive Analytics

- A 'general purpose analytical process that enables organisations to identify patterns in data that can be used to make predictions of various outcomes, not all of which have an impact on individuals' (OPCC, 2012, p.3).
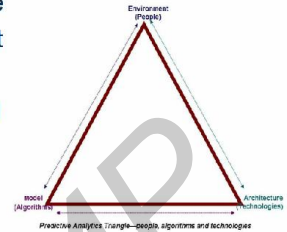
Conceptualising Predictive Analytics

- ... combines human skills and capability with technology such as machine learning of patterns in current and historical data and the application of algorithms not only to identify patterns in the data but also to forecast future probabilities of the outcome of those patterns (Ogunleye, 2014).

Conceptualising Predictive Analytics

- These widely accepted conceptions of predictive analytics have enabled data-driven organisations to take the sting out of the big data mayhem as well as strengthen their ability to 'generate better decisions, greater consistency, and lower costs' (CGI, 2013, p. 2).

Conceptualising Predictive Analytics

- What is clear from the previous slides is that people, tools and algorithms are critical in any predictive analytics project.

Limitations of predictive analytics: lack of subject knowledge

- A major consideration in the deployment of predictive analytics project is subject knowledge or the extent to which an organisation is familiar with the concepts of predictive analytics.

Quality of data

- As a statistical technique, data is central to predictive analytics.

- Poor attention to data quality issue could potentially lead to erroneous data mining and analysis results.

Model and modelling

- Model refers to a 'representation of the world, a rendering or description of reality, an attempt to relate one set of variables to another' (Miller, 2014, p. 2). Modelling, therefore, is a mathematical representation of an entity and very important in any predictive analytics project.

Model and modelling

- As Taylor (2012) explains, the modelling process has to be 'repeatable, industrial-scale' to ensure effective development of 'dozens or even thousands' of required predictive analytic models—in order to search for 'meaningful relationship among models and representing those relationships in models' (Miller, 2014, p.2).

Model and modelling

- Whatever the type predictive of model/s is deployed—be it regression or classification—user discretion, judgement and experience (or lack of them) will have impact on the outcome of predictive analytics.

Model and modelling

- A predictive modelling based on an out-of-date or incorrect data might 'wrongfully skew' analytics results (Viswanathan, 2013) or leads to wrong conclusions (Heger, 2014).

Model and modelling

- In deploying modelling in predictive analytics, there is sometimes a lack of a clear understanding of the difference between 'prediction' and 'projection' and how the two terms compare.

Model and modelling

- Analytics teams will run into trouble when projections are used to predict future trends when the underlying assumptions of their models are not *constant* or, at best, *susceptible* to seasonality trap.

Model and modelling

- In other words, when analytics teams are not clear about how the two terms compare, they might not be able to guarantee the *stability* in the phenomenon to be predicted (see also Elkan, 2013).

Model and modelling

- There is also a general tendency to equate correlation with causation in predictive analytics (Viswanathan, 2013)...
- ... especially if predictive analytics is conceptualised as involving 'searching for meaningful relationships among variables and representing those relationships in models' (Miller, 2014, p.2).

Return on investment

- ROI is an often-overlooked limitation in predictive analytics projects even though ROI is as higher as 250% in predictive analytics projects, according to a survey by the IDC (Vesset and Harries, 2011).

Return on investment

- There is also evidence that many organisations deploy predictive analytics projects and take little or no account of the ROI and those organisations that did have 'struggled to see a meaningful ROI (Accenture, 2013).
- The fact is that the levels of investment required for predictive analytics project is high (Harries, 2003; Accenture, 2013).

Legal and ethical

- Legal and ethical limitations exist especially where a company operates in different jurisdictions or cultures.
- The way information about customers are kept and mined and the 'extent to which data mining's outcomes are themselves ethical' with respect to individuals (Johnson, 2014; Kay, et al., 2012) can present challenges.

IT case for predictive analytics

- A limitation of predictive analytics is a perception that the case for a predictive project has to be about technologies or has to be predicated on information technology infrastructure.

Human factor

- Lack of consideration for human factor.

Human factor

- The introduction of predictive analytics technology will require attitudinal change and people in the organisation who have used to making decision based on intuition or gut feeling who consider themselves an essential part of the existing decision making process might feel that their toes are being stepped on.

Concluding remarks

- Predictive analytics is new approach to decision making as the technology enables organisations to make real time predictions with high degree of confidence.
- However, predictive analytics has its limitations – the main limitation is quality of data. Out-of-date or incorrect data can skew analytics results or produce wrong conclusions.

Concluding remarks

- Although, it is possible to use artificial intelligence based on machine learning algorithms to minimise the impact of data quality on the outcome of predictive analytics, but with the ever present danger of lurking variables or unknown factors, good algorithms might not be a *sustainable* alternative to good data quality assurance.

References

- Accenture (2013) Analytics in Action: Breakthroughs and Barriers on the Journey to ROI, online: http://www.accenture.com/sitecollectiondocuments/pdf/accenture-analytics-in-action-survey.pdf; accessed: 28.5.15.
- CGI (2013) Predictive analytics. The rise and value of predictive analytics in enterprise decision making, White paper, [Online]. www.cgi.com; accessed: 20.4.2015.
- Elkan, C. (2013) 'Predictive analytics and data mining', Online: http://cseweb.ucsd.edu/~elkan/255/dm.pdf; accessed: 10/3/15.
- Harries, D. H. (2003) Predictive Analytics and ROI: Lessons from IDC's Financial Impact Study, International Data Corporation (IDC), Online: http://www.analyticalinsights.com/web_images/IDC-PredictiveanalyticsandROI.pdf; accessed: 3.5.15.
- Vesset, D. and Harries, D. H. (2011) 'The Business Value of Predictive Analytics', White Paper, International Data Corporation (IDC), Online: http://www.nexdimension.net/resources/products/ibm/spss/ibm-spss-predictive-analytics-business-value-whitepaper.pdf; accessed: 3.5.15.

References

- Heger, D. (2014) Big Data Analytics—'Where to go from Here', *International Journal of Develkopments in Big Data and Analytics*, 1, 1, pp.42-58.
- McCann, D. (2014) Predictive Analytics: How Clear Is the ROI?, CFO Magazine, Online: http://ww2.cfo.com/technology/2014/07/predictive-analytics-clear-roi/; accessed: 26.5.15.
- Miller, M. T. (2014) Modeling Techniques in Predictive Analytics: Business Problems and Solutions with R, Pearson Education, Inc.
- Ogunleye, J. (2014) The Concepts of Predictive Analytics, *International Journal of Developments in Big Data and Analytics*, 1, 1, pp. 86-94.
- OPCC (2012) The Age of Predictive Analytics: From Patterns to Predictions. Report prepared
- by the Research Group of the Office of the Privacy Commissioner of Canada, available: https://www.priv.gc.ca/information/research-recherche/2012/pa_201208_e.pdf; accessed: 23 July2014.
- SAS Institute (2012) Big Data Meets Big Data Analytics, White Paper. SAS Institute Inc. Available:http://www.sas.com/content/dam/SAS/en_us/doc/whitepaper1/big-data-meets-bigdata-analytics-105777.pdf; accessed: 2 January 2014.

References

- .Viswanathan, V. (2013) 'Avoiding Pitfalls in Predictive Analytics Models', Online: http://data-informed.com/avoiding-pitfalls-in-predictive-analytics-models/; accessed: 20.11.14