

Can big data studies know your thoughts and predict who will win an election?

August 9 2016, by David Tuffley And Bela Stantic



Who'll be the next president of the United States – Donald Trump or Hillary Clinton? Credit: Flickr/torbakhopper, CC BY-NC-ND

Who will win the battle for the White House to become the next President of the United States is a topic of hot debate.

Much of that debate is taking place online, with plenty of people blogging, tweeting or updating social media with their thoughts on Donald Trump versus Hillary Clinton.

This provides us with a rich source of information about what people are thinking and feeling about the election race.

When our Big Data and Smart Analytics Lab analysed those comments on Twitter towards the end of July, it predicted that if the US Presidential election had been held at that time, Trump would have been the winner over Clinton. Two weeks later, a swing towards Clinton is making the result less certain.

But how accurate is this type of prediction? Well, the same lab using the same method predicted and announced in a public lecture a week before the Australian federal election that the Coalition would win over the ALP.

So it will be interesting to see what the lab's prediction says of Trump versus Clinton as we get closer to the US election in November.

Volume of traffic

The amount of data everyone on the internet generates in [just a single second](#) is truly staggering, and it is continuing to grow, with [estimates that annual global traffic](#) will pass the zetabyte barrier by the end of 2016. That's one trillion gigabytes (GB) of data, roughly 134GB for every man, woman and child on the planet!

Big data analytics is the discipline faced with the challenge of managing the sheer volume of data and turning it into something useful.

It makes predictions about the future based on the patterns of the past, and finds relationships buried in the data that no one has noticed. It is also very useful for running simulations to see the consequences of a particular course of action.

Its predictive power is also increasing. The analytics use smarter, faster algorithms to perform [deep learning](#) on the large volumes of data drawn from diverse sources, including from publicly accessible social media.

Surprisingly accurate predictions can now be made about trends in financial markets, fashion and tourism to name a few areas, as well as the outcome of elections.

It's already helping businesses and governments to make strategic, well-informed decisions across a wide range of applications.

The human approach

We can improve the accuracy of simple big data prediction by using what we call the human sensor approach, developed at the Smart Analytics Lab. This recognises the whole person as an excellent data collection source.

Oceans of real-time data from social media, discussion groups, blogs and review websites are brought together along with other freely available [open source](#) data.

In the field of open data and data analytics there is a legally enforceable code that all data, regardless of where it is from, is properly anonymised. Anything that could relate to individual people is removed.

A sentiment analysis is then performed on the data that reveals what people are actually thinking about something, including how they feel about it.

This holistic approach manages to remedy many of the previous accuracy problems. For example, micro-blogs such as Twitter are very short, full of spelling errors and slang language. The new approach is

able to deal with this problem.

In a project sponsored by the [National Environmental Science Program](#) the Smart Analytics Lab is looking to use this type of data analysis to provide information on the Great Barrier Reef.

By sampling different sources of data, including [social media](#) postings and photos, it's possible to recognise types of fish, how plentiful or scarce they are and the extent of coral bleaching. Integrating all of this data and applying deep learning allows potential environmental issues on the Reef to be identified early.

On the Gold Coast, a similar funded project is underway to gauge the satisfaction or otherwise of visitors to the Coast. Doing a sentiment analysis on the data tells us how people really feel about their holiday experience. For example, what was the accommodation like, were the beaches enjoyable and how about overall value for money?

This is information that can be used strategically by tourism operators to hone in on where they need to improve and what they already do well and could be doing more of. It can also tell the city government similar things about the public infrastructure.

Looking for a new job?

These advances in [big data](#) analytics are made possible by a convergence of three factors; cheaper yet more powerful computers, more publicly available data and smarter algorithms that learn over the time from repeated cycles of analysis.

But people skilled in data science [are in high demand](#) and are becoming increasingly rare in the employment market.

This is partly because in addition to skills in data management and machine learning the work calls for skills such as behavioural psychology, bio-infomatics, business, social anthropology and linguistics. These are skills not usually associated with the IT world.

Progressive universities are catching up with demand for data science programs that go some way to meeting the growing need. With so much talk of the jobs that won't exist in the future, it's good to know that there are emerging sectors where people will still be needed, for now at least.

This story is published courtesy of [The Conversation](#) (under Creative Commons-Attribution/No derivatives).

Source: The Conversation

Citation: Can big data studies know your thoughts and predict who will win an election? (2016, August 9) retrieved 26 April 2024 from <https://phys.org/news/2016-08-big-thoughts-election.html>

| |
|--|
| <p>This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.</p> |
|--|