Can big data predict the outcome of major political events?

Written by

Julio Amador

Published

25 February 2017

Category

Technology

Key topics

Data, Social Media

Dr Julio Amador, Research Associate at Imperial Business School, looks at how big data was used to monitor people's behaviour online in reaction to Brexit and the US election

Big data is transforming society. Not so long ago, this phenomenon was confined to within the walls of universities, but now big data plays a key role across most industry sectors – affecting how companies function, attract new customers and streamline their business strategy.

The data revolution is changing the way businesses operate and compete in the global market and has transformed a number of industries including healthcare, manufacturing, education and agriculture.

At the KPMG Data Observatory, housed within the Data Science Institute at Imperial London, researchers can now create visualisations of global business and social trends on everything from cryptocurrency transactions to migration patterns, on a

bank of screens powered by 32 computers.

The aim of the data observatory is for researchers across Imperial London and Imperial Business School to collaborate with businesses to analyse and visualise big data to generate insights that can help businesses improve their products and services. Researchers also hope to use the observatory to better understand subjects such as online fraud, climate change and the solar system.

But the observatory can also be used for analysing people's behaviour online in the lead up to a major public vote. We looked at this phenomenon in two recent studies where we looked at how people used Twitter on the eve of Britain's decision to leave the European Union and on the night of the US election.

In the first study on Brexit, conducted by researchers at Imperial Business School, we found that Twitter users who sent messages to both those who shared their views and others with opposing viewpoints, were more likely to spread their message further online. The research showed that Twitter users who supported the Leave campaign were three times more likely to engage with people from both sides of the argument online than people who backed Remain.

The most active users in the Leave group were shown to both engage in dialogue and provide information whereas people supporting Remain were more likely to just provide information to fellow users. Our research team tracked 21,000 messages sent from UK Twitter users who were engaged in discussion about the EU Referendum using the following hashtags: #Brexit, #LeaveEU, #Go, #Remain, #Strongerin, #Voteremain, #bremain and #voteleave.

We tracked both the topics in tweets to other users and the overall volume of Twitter traffic in relation to the EU Referendum across the UK, to get an insight into how polarised people's views are in the lead up to a major political event.

Our analysis showed that London was the region with the highest amount of Twitter engagement on Thursday 9 June 2016, with a huge surge in online discussion between 8.00-10.00pm during an ITV televised debate.

In our second study, we found that Twitter users who backed Donald Trump were also more likely to reach out to others with opposing viewpoints.

For this study, we analysed 3.5 million tweets made in the 12 hours leading up to the election result on polling day. We found that on Twitter, supporters of Mr Trump were more likely to engage with others of opposing political views to spread their message further.

Using data analysis techniques, we compared the volume and sentiment of tweets made by both Republican and Democrat supporters using the following hashtags: #MakeAmericaGreatAgain, #I'mwithher, #Election2016 and #Electionnight.

We found that in the 12 hours before the polls closed there was a surge of 3,000 tweets using the hashtag #MakeAmericaGreatAgain (used by Mr Trump supporters) compared with only 1,400 for #I'mwithher (used by supporters of Hillary Clinton). We concluded that Trump supporters made greater efforts to win over voters on social media than Clinton supporters, right up until the polls closed.

We also found a higher level of Twitter conversation about the election in swing states such as Florida and Pennsylvania.

We can draw parallels between the level of engagement on Twitter in terms of volume and sentiment of tweets coming from Democrat supporters, to the level of Twitter engagement put out by the Remain campaign on the day of the EU Referendum vote. In my view, these results show that supporters of Remain and Hillary Clinton could have done more to reach out to voters online.

The US election was also the inspiration for a separate study from my colleague Dr Miguel Molina-Solana, a Research Associate based at the Data Science Institute. Miguel's team analysed the content of Mr Trump's and Mrs Clinton's speeches in the lead up to the US election to see which topics the presidential candidates touched on the most.

Using data analysis, the researchers examined transcripts of both Mr Trump's and Mrs Clinton's speeches, breaking the content down into a matrix featuring key words and topics to understand which topics each candidate spoke more about.

The results show that Mr Trump's speeches were more likely to be emotive and focus on topics such as leisure, family, religion, music, 'positive feelings' and America. While both candidates spoke about law and order, our research implies that Mr Trump was better able to deliver his message on law and Government through fewer speeches than Mrs Clinton.

So we've seen how big data can provide fascinating insights into people's behaviour online in reaction to a major public vote. But, in future – could we use big data to predict the outcome of these events before they happen?

In a separate study we did with the London School of Economics, we used machine learning and artificial intelligence to separate the tweets from the Twitter/Brexit study into Leave and Remain groups. Our findings indicate that Twitter data was able to predict online support for the Leave and Remain campaigns as accurately as more general internet polls.

Written by

Julio Amador

Published

25 February 2017

Category

Technology

Key topics

Data, Social Media

Share

About Julio Amador

Junior Research Fellow
Dr Julio Amador is currently dedicated to studying advertising in social networks, microfinance and data-mining algorithms.

Read Imperial Professional Web Page for more information and publications.

Monthly newsletter

Receive the latest insights from Imperial Business School

Sign up now

Opaque data practices keep Big Tech uncompetitive

Opaque data practices are leading to a lack of skills, ideas and thinking, and eroding consumer data protection within Big Tech companies.

Read more