Universidad de Granada

Researchers model a system based on artificial intelligence that enables election results to be fore

03/11/2020

Research news

[Spanish version]

Scientists from the University of Granada have applied Artificial Intelligence techniques to the analysis of huge volumes of data from Twitter, during the previous US election campaign to create a political forecasting system

Researchers from the Department of Computer Science and Artificial Intelligence at the University of Granada (UGR) have modelled a system based on artificial intelligence techniques that enable election results to be forecast by analysing opinions on Twitter.

In a study published in the international journal IEEE Access, the UGR scientists explain their descriptive Big Data system capable of handling huge volumes of unstructured information (in the form of a 'data lake') derived from Twitter. Using this approach, they were able to create a political forecasting system and validate it with the real-life 2016 US elections, in which Donald Trump won against Hillary Clinton.

Political talk is perhaps more prevalent than ever before—one need only look tosocial networks for evidence of this, and the sheer amount of posts and threadsdevoted to political topics each day. One of the most widely used social networks forthese purposes is Twitter, where the opinions of parties, leaders, and activistscombine with those of people simply interested in politics. The ability to effectivelyprocess this data and convert it into knowledge is a laborious task that deliversbenefits for innumerable fields, from academia to business or journalism.

The UGR study is the result of an endeavour to 'summarize' a large volume of data and reduce it to clear, concise information that can contribute value to a research query. The system in question was developed by José Ángel Díaz García, María Dolores Ruiz and María José Martín-Bautista from the UGR's Department of Computer Science and Artificial Intelligence. It was tested on a real-life comparative problem concerned with two politicians and their respective policies: that of Donald Trump and Hillary Clinton, in their head-to-head clash in the November 2016 US general elections.

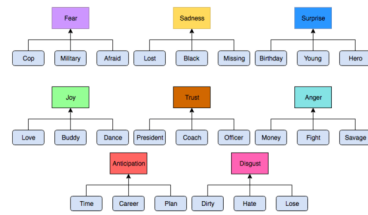## Analysis of sentiments and emotions

The system devised by the UGR scientists provides a series of associations between concepts and discussions on Twitter about the two politicians—in a format that is easy to interpret and explain—together with the sentiments and emotions generated by these debates.

"At the heart of our system are what we call unsupervised artificial intelligence techniques—that is, techniques that do not rely on databases having been pre-labelled in order to be trained and used," the authors explain.

Among these techniques, of particular importance are 'association rules', as these enable sentiment analysis to be conducted by means of sentiment lexicons and dictionaries. "Today, these techniques are of enormous value because they provide readily interpretable and easily understandable solutions. They enable straightforward data traceability and provide easily-explained results that may be used by people with no technical knowledge, thus democratizing access to artificial intelligence," the authors continue.

This new descriptive approach differs from the traditional 'machine learning' models geared to predictive sentiment analysis. Those require large pre-labelled databases (very hard to achieve in relation to social networks, due to the volatility of the topics concerned), and typically offer solutions that are extremely difficult to interpret due to the highly complex mathematical adaptations.

Analysis of the results achieved by the new system endorses its capacity to obtain

association rules and sentiment patterns with significant descriptive value in the case of its application to the US elections. Thus, parallels between these patterns and real-life events can be drawn.

Some of the parallels discovered by the system may be those, for instance, that establish a very strong link between the words prohibition/service/transgender and Donald Trump. This shows that the current US president was linked to transgender people being banned from Military Service—a move that was already being considered in 2016 and was confirmed in 2017.

Regarding sentiments, the system reveals that there was a higher level of anger in US society directed toward Hillary Clinton than toward Trump. The latter, by contrast, stood out for his association with the emotion of 'trust'—in other words, the Tweets posted about Trump were from people with a high degree of confidence in him as President.

If we take into account that the data were processed during the electoral campaign, a parallel could therefore even be drawn in the subsequent results that led Donald Trump to victory.

Bibliography:

J. A. Diaz-Garcia, M. D. Ruiz and M. J. Martin-Bautista (2020), 'Non-Query-Based Pattern Mining and Sentiment Analysis for Massive Microblogging Online Texts', IEEE Access 8: 78166-78182. DOI: 10.1109/ACCESS.2020.2990461.

Image captions:

1. Words associated with emotions

2. Generalization of words (by sentiment)

3. The authors of this research. From left to right: Maria Jose Martin-Bautista, J. Angel Díaz-García and María Dolores Ruiz

Media enquiries:

José Ángel Díaz García

Department of Computer Science and Artificial Intelligence, University of Granada

Email: @email