

The creative data interpreters

August 29 2017, by Meryem Riahi



The software uses machine learning to create so-called word clouds, thus visualising the linking of words in, for example, tweets. Credit: SpinningBytes

The ETH spin-off SpinningBytes programs software that uses machine learning not only to analyse but also to understand huge amounts of data. It enables customised solutions to be developed for numerous IT problems, and allows new insights to be gained from previously unused data.

It all began a few years ago: Mark Cieliebak, Martin Jaggi and Fatih Uzdilli were computer science researchers at ETH and Zurich University of Applied Sciences (ZHAW), and published their technologies in scientific papers. However, they did not gain much publicity. In order to change that, they founded the ETH spin-off SpinningBytes in 2015 and

made their programs freely available on their homepage.

The spin-off now offers primarily project-related programming of data science software alongside complete technology solutions. "We develop programs that can analyse data and to a certain extent understand it," explains CEO Cieliebak. In order to build the software, the spin-off needs data on which to base the program and from which it can learn: "Our software looks at the existing data, produces statistics about certain regularities and generates new knowledge, which then guides it."

Analysing and classifying articles

One example is the classification and categorisation of huge amounts of text – such as the software that SpinningBytes has programmed for the Swiss Economic Archive. The archive has been collecting reports on the Swiss economy since 1890, all of which are categorised according to the same pattern. Previously, archivists examined every text, but the program can now take over part of this work. The software analysed 30,000 categorised articles and learned the assignment rules.

Digital customer service

The computer scientists at SpinningBytes not only deal with written texts, they also develop programs that can recognise and understand the human voice – and give an answer. "There is immense potential for such dialogue systems, particularly in customer service," says Cieliebak, "because a service hotline will often run through the same forms of dialogue." These repetitive conversations could be automated using machine learning. In future, for example, a health insurance company could settle the first standardised contact with potential customers using a digital form. "The software will not replace people, however," emphasises Cieliebak. As soon as the conversation deviates too much

from the standardised phrases, the software can no longer resort to automated responses and has to pass the customer on to a [customer service representative](#).

Calculating heart attack risk using Twitter

Prognoses can also be generated by analysing huge amounts of data. In a new project, one of SpinningBytes' programs uses tweets to investigate the risk of heart attack in different regions. So how does it work?

"Among other things, a heart attack has to do with whether you are happy or not," says Cieliebak, and explains further: "The language in the tweets allows conclusions to be drawn about levels of satisfaction, and by linking this to other statistical data, statements can be made about the [heart attack risk](#) in a particular area." Other such prognostic programs are currently in the planning stage.

Provided by ETH Zurich

Citation: The creative data interpreters (2017, August 29) retrieved 23 April 2024 from <https://techxplore.com/news/2017-08-creative.html>

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.