Publication: MIT Technology Review

Date: 22 February 2018

Headline: Al is learning how to spot risky websites for you

Al is learning how to spot risky websites for you

Machine learning can sniff out tell-tale signs of shady URLs so you don't get phished.

The problem: The internet is riddled with websites set up for the sole purpose of stealing a user's information or installing malware on a victim's machine. Antivirus companies blacklist them as fast as they can, but with new sites launched every day, it's a Sisyphean effort to keep up.

Al to the rescue: A new system called URLNet uses neural networks that look at character-level and word-level combinations in—you guessed it—the site's URL to detect a site's risk. URLs contain clues to whether a site is malicious, like length and misspelled domain names.

Results: The researchers trained URLNet on two data sets, one containing a million legit and malicious URLs and one with five million. In each case, URLNet beat other current systems at detecting suspicious sites.

Source: www.technologyreview.com/the-download/