



Breaking Google's reCAPTCHA v3

Any person who has spent some time on the Internet has encountered CAPTCHAs meant to deter bots from web pages by requiring a problem to be solved, which only humans are supposed to be able to do. However, in recent years, machine learning models have become more advanced and have been shown to break CAPTCHAs, which begs the question: Are CAPTCHAs even relevant anymore? Methods to break CAPTCHAs involve, for instance, sending the problem to a third party and having a human solve it for a small fee. In general, most websites use Google's reCAPTCHA v2 and v3. The v2 CAPTCHA asks the user to solve an image classification or segmentation task. The v3 CAPTCHA is more subtle, and the user does not interact with it visibly. In prior work, we have shown to be able to break reCAPTCHA v2, but now we want to break v3 as well using, for instance, reinforcement learning.

This project will focus on determining what features the reCAPTCHA system uses to determine if a user is a bot or not. This also entails determining how much each feature impacts the classification.

Requirements:

Prior experience and a strong interest in machine learning, reinforcement learning, and web development are recommended. Creativity and programming skills are advantageous.

Interested? Please contact us for more details!

Contact

- Andreas Plesner: aplesner@ethz.ch, ETZ G95