Defeating Google: CAPTCHA Edition

By

John Haag, Zach Kovalenko, and Thomas Shear

In this Artificial intelligence project, the three group members have decided to create an artificial intelligence program that will defeat Google’s CAPTCHA. CAPTCHA or RECAPTCHA is a test that Google has created similar to a Turing test to determine whether a user is human or machine. The Turing test was conceptualized by the father of Computer Science Alan Turing. At the time of conception, the test was known as the Imitation Game. This test is one of the fundamental principles of Artificial intelligence, and is the idea that a machine has the ability to exhibit intelligent behavior that is equal to or indistinguishable from a human. This concept is important when dealing with CAPTCHA and defeating it. Captcha is security test that is held on a wide variety of websites that keep robots from entering certain areas of a site. These areas tend to be high network traffic areas like the checkout in most E-commerce websites. As most of our group members buy and sell things on the internet, solving a way to defeat Captcha with a Python bot is important to our profit margins. If you are unfamiliar with how captcha works here is a brief overview. Once the user finds a site that has Captcha the test asks the user a few questions to verify it is a human, but instead of using simple HTML, CSS or JavaScript, that a computer could read in plain text or one of those languages Captcha uses pictures, audio or no questions at all to figure out if the user is a human or robot. This is where the challenge of Artificial intelligence comes into play, we plan to write a program using the Python programming language, and other artificial intelligence techniques to defeat Captcha. At this time it is unknown what method of Artificial intelligence techniques we will use to solve this problem.