Text-based CAPTCHA Using Python

GitHub Link For The Code of Project: -

https://github.com/chirag-m1/INT213-PYTHON-PROJECT.git

Project Report

by

Chirag Monga

Registration Number – 11904816

Section – K19KK

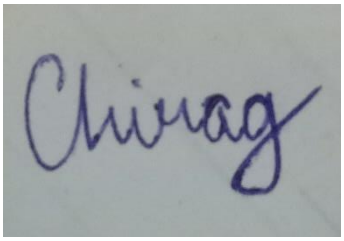
Roll Number – 21

Department of Intelligent Systems
School of Computer Science Engineering
Lovely Professional University, Jalandhar

October 2020

Student Declaration

This is to declare that this report has been written by us. No part of the report is copied from other sources. All information included from other sources have been duly acknowledged. We aver that if any part of the report is found to be copied, we shall take full responsibility for it.



Chirag Monga

11904816

Place – Lovely Professional University
Date – November 01, 2020

BONAFIDE CERTIFICATE

Certified that this project report titled “Text-based CAPTCHA using Python” is the bonafide work of Mr. Chirag Monga (Reg. No: 11904816) who carried out the project work under my supervision.

……

Ankita Wadhawan,
ID: - 23891,
Assistant Professor,
School of Computer Science and Engineering

TABLE OF CONTENTS

INTRODUCTION

A CAPTCHA is a program that protects websites against bots by generating and grading tests that humans can pass but current computer programs cannot. For example, humans can read distorted text as the one shown below, but current computer programs can't. These distorted texts are known as Text-based CAPTCHA.

BACKGROUND

The term was coined in 2003 by Luis von Ahn, Manuel Blum, Nicholas J. Hopper, and John Langford. The most common type of CAPTCHA (displayed as Version 1.0) was first invented in 1997 by two groups working in parallel. This form of CAPTCHA requires someone to correctly evaluate and enter a sequence of letters or numbers perceptible in a distorted image displayed on their screen. Because the test is administered by a computer, a CAPTCHA is described as a reverse Turing test.

This user identification procedure has received many criticisms, especially from people with disabilities, but also from other people who feel that their everyday work is slowed down by distorted words that are difficult to read. It takes the average person approximately 10 seconds to solve a typical CAPTCHA.

OBJECTIVE

The purpose of CAPTCHA

The basic purpose of CAPTCHA is to distinguish between a machine and a human being. As a machine or robot can be used for filling forms with random data but a image or audio to be entered cannot be done by a robot or machine and hence a human only can do it.

- Block spammers and bots that try to automatically get email addresses or try to automatically sign up for or make use of Web sites, blogs or forums.
- To test that the end user is machine or human.
- This is beneficial in stopping many types of bots or crawlers are programmed that can be used to autofill the form again and again by which dos attack can be performed in the sense of memory consumption.

APPLICATIONS

CAPTCHAs have several applications for practical security, including (but not limited to):

• Preventing Comment Spam in Blogs. Most bloggers are familiar with programs that submit bogus comments, usually for the purpose of raising search engine ranks of some website (e.g., "buy penny stocks here"). This is called comment spam. By using a CAPTCHA, only humans can enter comments on a blog. There is no need to make users sign up before they enter a comment, and no legitimate comments are ever lost!

• Protecting Website Registration. Several companies (Yahoo!, Microsoft, etc.) offer free email services. Up until a few years ago, most of these services suffered from a specific type of attack: "bots" that would sign up for thousands of email accounts every minute. The solution to this problem was to use CAPTCHAs to ensure that only humans obtain free accounts.

• Protecting Email Addresses From Scrapers. Spammers crawl the Web in search of email addresses posted in clear text. CAPTCHAs provide an effective mechanism to hide your email address from Web scrapers. The idea is to require users to solve a CAPTCHA before showing your email address.

• Online Polls. Many websites release online poll asking some useful questions or quizzes which they require

honest and unbiased answers (a dangerous question to ask over the web!). As is the case with most online polls, IP addresses of voters were recorded in order to prevent single users from voting more than once. Can the result of any online poll be trusted? Not unless the poll ensures that only humans can vote.

• Preventing Dictionary Attacks. CAPTCHAs can also be used to prevent dictionary attacks in password systems. The idea is simple: prevent a computer from being able to iterate through the entire space of passwords by requiring it to solve a CAPTCHA after a certain number of unsuccessful logins. This is better than the classic approach of locking an account after a sequence of unsuccessful logins, since doing so allows an attacker to lock accounts at will.

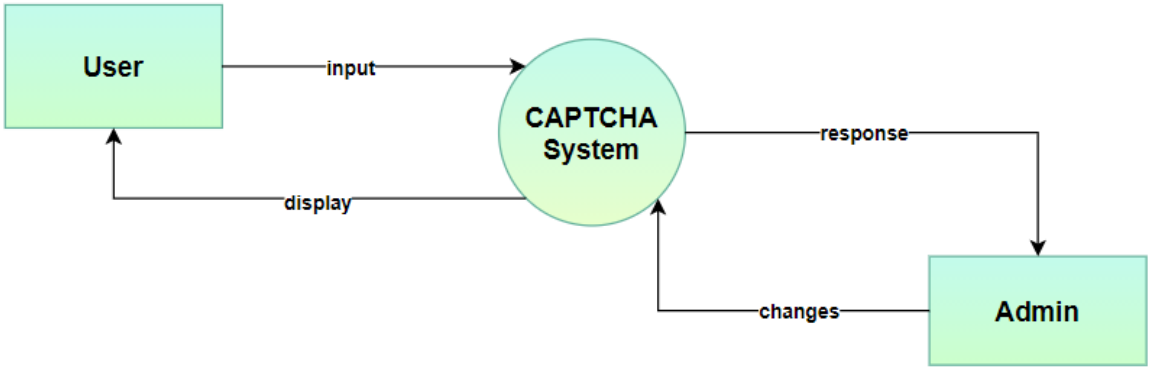
• Search Engine Bots. It is sometimes desirable to keep webpages unindexed to prevent others from finding them easily. There is an html tag to prevent search engine bots from reading web pages. The tag, however, doesn't guarantee that bots won't read a web page; it only serves to say "no bots, please." Search engine bots, since they usually belong to large companies, respect web pages that don't want to allow them in. However, in order to

truly guarantee that bots won't enter a web site, CAPTCHAs are needed.

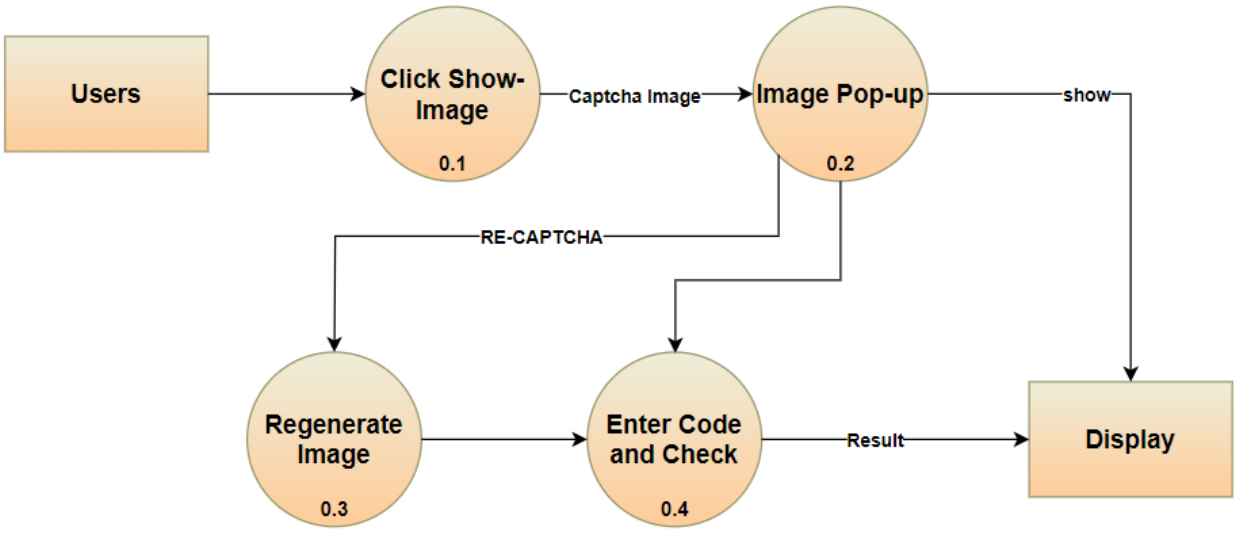
Worms and Spam. CAPTCHAs also offer a plausible solution against email worms and spam: "I will only accept an email if I know there is a human behind the other computer." A few companies are already marketing this idea.

DESCRIPTION

Level 0 DFD

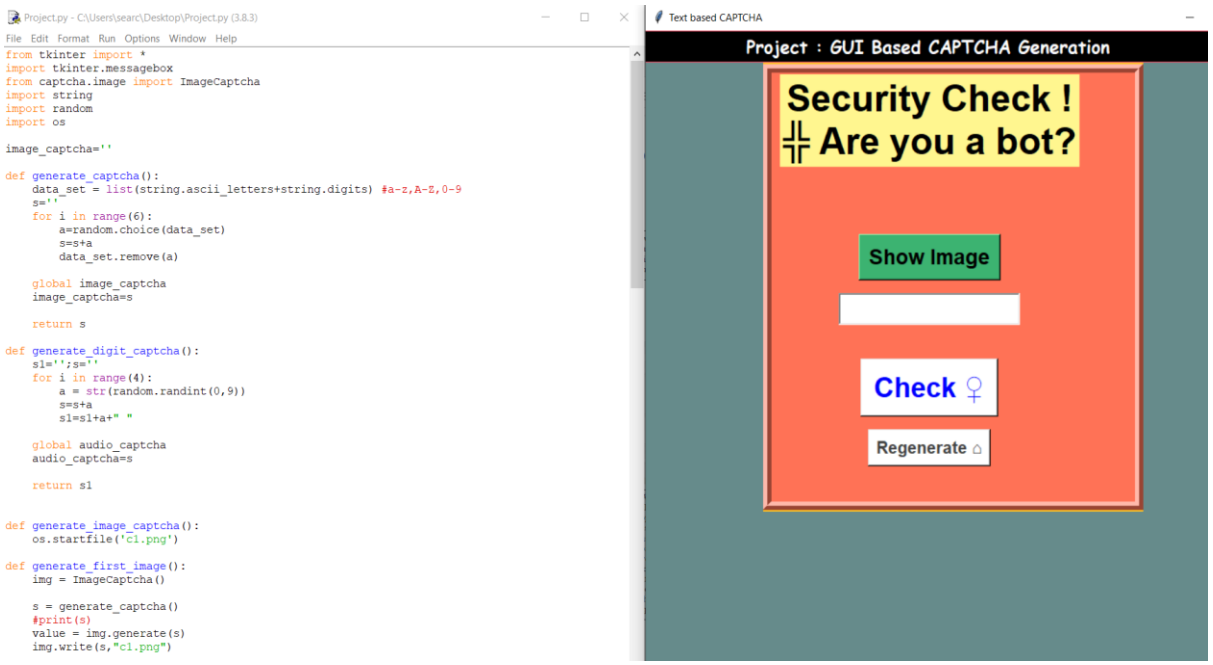


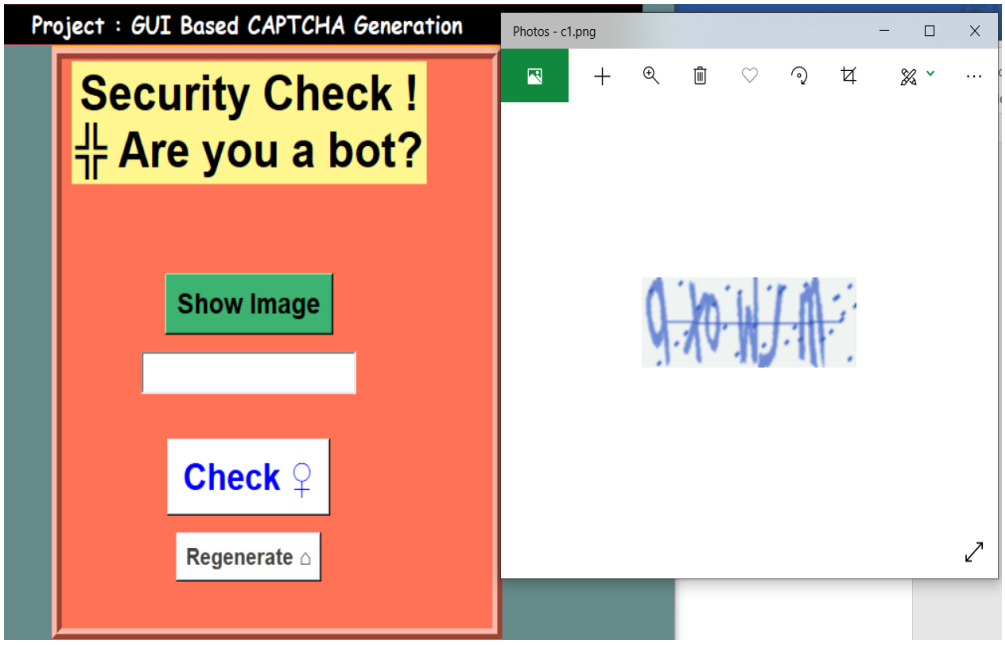
Level 1 DFD



IMPLEMENTATION OF PROJECT

When we execute the program, a GUI based interface appears.

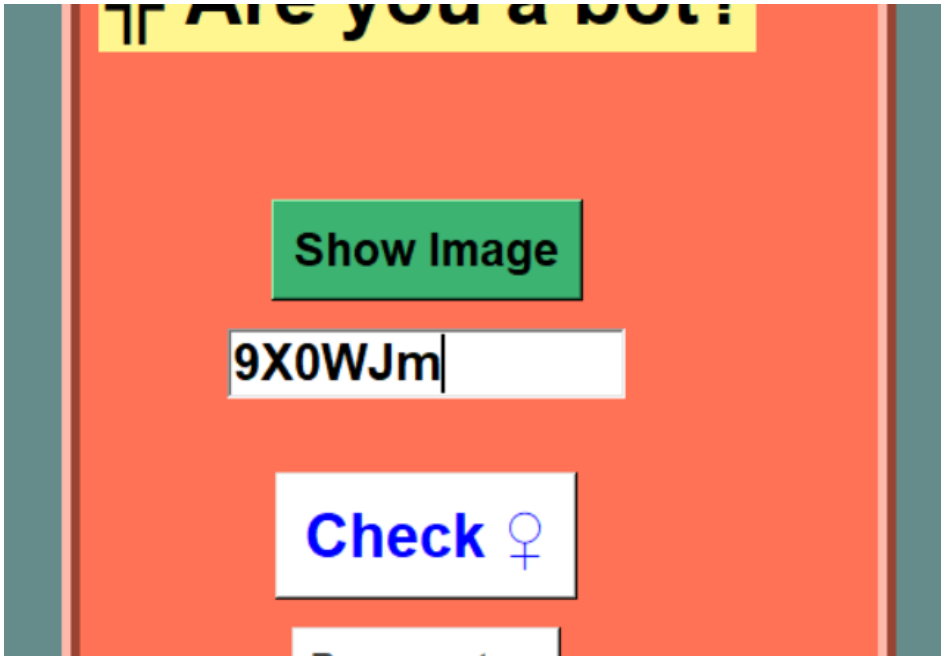


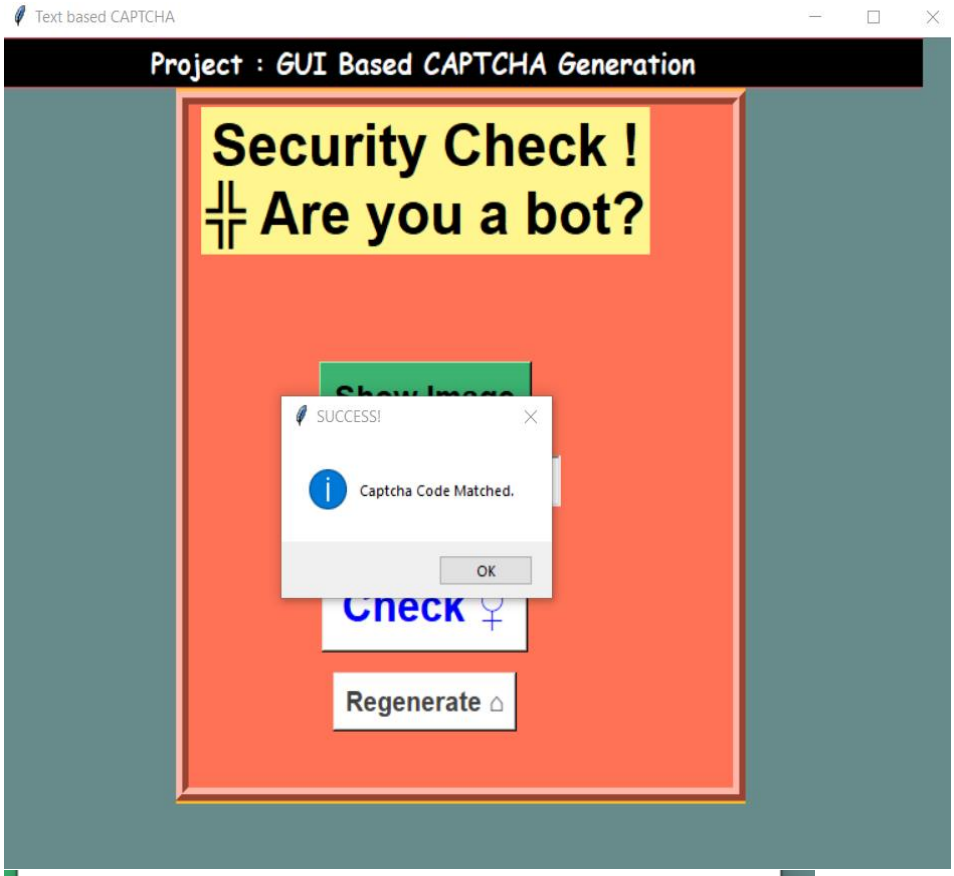


On clicking 'Show Image' button, a CAPTCHA code image will pop-up on the screen.

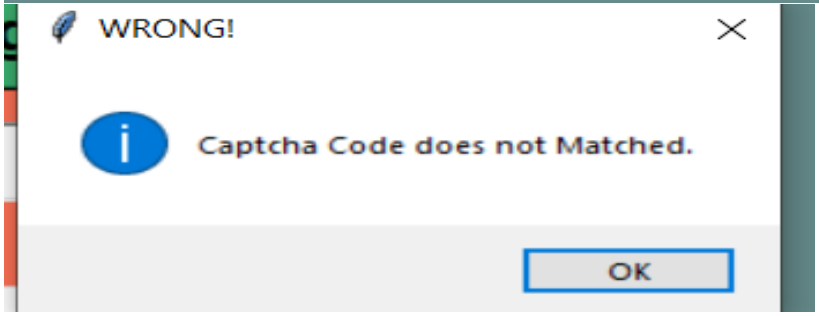
Code contains random distorted alphabets and numbers.

Enter the CAPTCHA displayed in the blank white box and click ‘Check’ button





If the entered CAPTCHA code is correct, a message box showing “Captcha Code Matched” appears.



If the entered CAPTCHA code is incorrect, a message box showing “Captcha Code does not Matched” appears.

To generate a new CAPTCHA, press ‘Regenerate’ button following which a new image will pop-up showing different CAPTCHA code.

OUTCOME

The result of the assignment results out the following outcomes:

- It helped us to understand the benefits of CAPTCHA and necessary knowledge about it.
- We're able to understand the different types of CAPTCHAs.
- It made us understood the method to use the GUI programming into much detail.
- It helped us to know more about the importance GUI application in our daily life.

TECHNOLOGY AND PLATFORM USED

Technology used in PYTHON**:**

- Tkinter to build up GUI
- String, Random and Captcha modules in Python.
- Message Box to display output message
- Conditional statements for conditions

Platform used:

- IDLE
- Jupyter Notebook