Idea Introduction

Protectnetic - Child Abuse Detection (Website)

Tech Stack

1. Python / Django
2. HTML
3. JS/CSS
4. Bootstrap

Resources Used

1. OpenCV
2. TensorFlow
3. Beautiful Soup
4. Selenium
5. Tor-socks (for proxy)
6. Requests (to scrape websites)

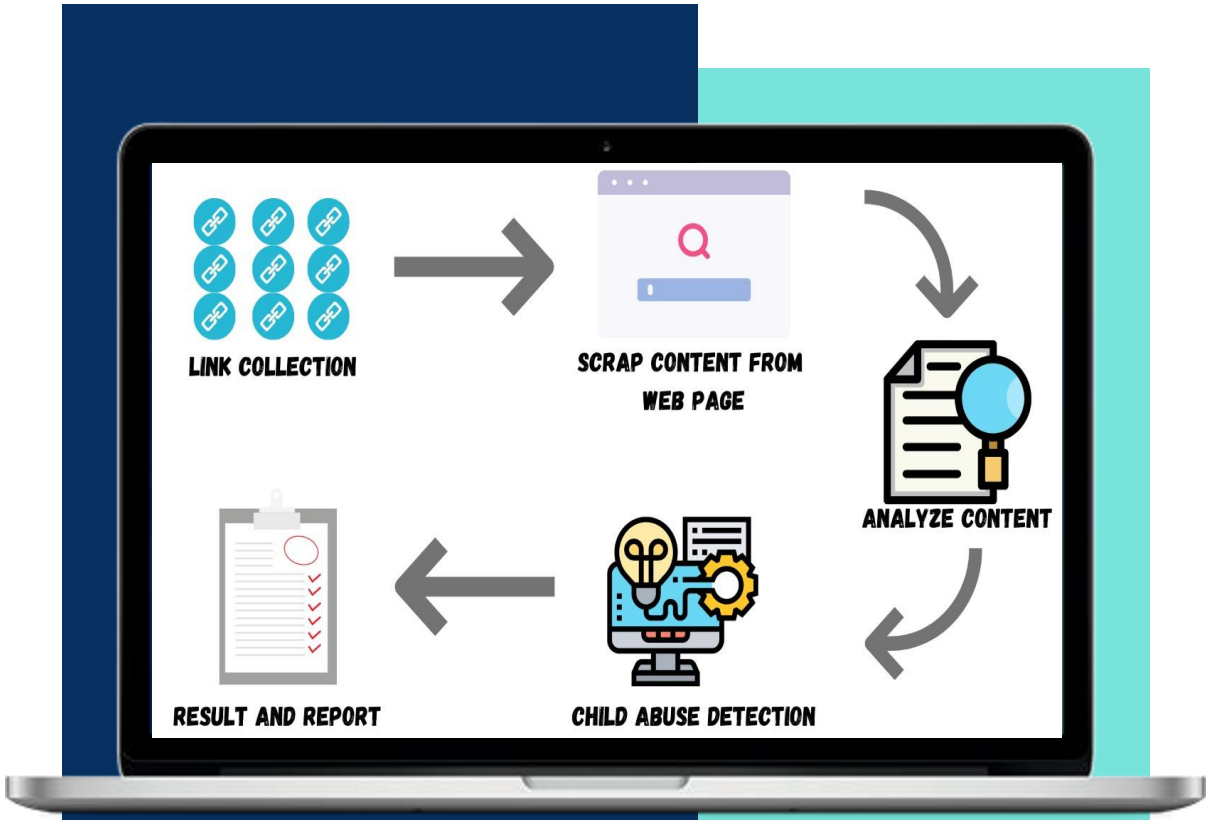
Any Third Party API/Services used

1. Currently None

IDEA-OUTCOME – HEADLINE-

To make a website for scraping the dark web website, analyse the content in it and generate a report with the required details throw an alert if child abusive content is found on the provided dark web website.

Your Approach Towards Idea



Approach

To scrape the dark web website, analyse the content in it and generate a report for the required details& throw an alert if child abusive content is found on the provided dark web website.

We targeted the problem by gathering content about the given domain & analysing it's contents & proposed the result based on our research.

We trained a ML model to detect child abusive content based on scrapped images from the dark web.

Future Enhancements:

To Train a ML model which detects child abusive content based on scrapped html.

Development Pipeline

Proposed Development

1. Collect Links
2. Scrape Links
3. Train ML Model to detect age & Sexual Content.
4. Put it all together
5. Make Website

Step A

Documentation

1. ReadME.md
2. Installation Docs.
3. Help Docs.

Step B

Our Repository Links and References:

1. Git Repo
2. Diggy Spidey

Step C

Any Third party API used :
Currently None

Step D

Proposed UI and functional flow

Flow :
Home Page -> Search Page (Enter any dark web link) -> Result Page -> Report and Results

Step E

Vision of Innovation/Idea/Solution

Explain How you developed Idea ?

By researching about dark web & problem statement we figured out what the problem was & brainstormed what can be the possible outcomes to the said problem.

How your idea is different and innovative form other ideas ?

Our Idea involves Image Processing to detect Abusive Content, also we are planning to use NLP to perform sentimental analysis on the scraped data, while other ideas might include different approaches to solve the same problem.

How much time it will take in conversion as a final product ?

– 2.5 Months Tentatively

Any early stage innovation detected while developing the solution ?

An NLP model can used to directly detect CAC (Child Abusive Content) from the HTML of the Dark-Web Link. This not only the optomise the solution but also reduce the time for CAC detection.

Thank you for your time!