

## A

**Project Report On**

## "Phishing Site Detection using Machine Learning"

(SOFTWARE GROUP PROJECT-4)

### Prepared by

VAIBHAV KANPARIYA (19DCS051) SHREY GAJJAR(19DCS034)

### Under the Supervision of

### Prof Shivangi Mehta

### Submitted to

Charotar University of Science & Technology (CHARUSAT) for the Partial Fulfillment of the Requirements for the Degree of Bachelor of Technology (B.Tech.)

in Computer Science & Engineering (CSE)

for 6th semester B.Tech

**Submitted at**

**Devang Patel Institute Of Advanced Technology And Research**

**Faculty of Technology & Engineering CHARUSAT At: Changa, Dist: Anand, Pin: 388421.**

**November , 2021**

### DECLARATION BY THE CANDIDATE

I hereby declare that the project report entitled “A Phishing Site Detection using Machine Learning” submitted by me to **Devang Patel Institute Of Advanced Technology And Research, Changa** in partial fulfilment of the requirement for the award of the degree of **B.Tech** in Computer Science And Engineering, from Devang Patel Institute Of Advanced Technology And Research /FTE, is a record of bonafide Software Group Project-4 (project work) carried out by me under the guidance of **Prof. Shivangi Mehta**. I further declare that the work carried out and documented in this project report has not been submitted anywhere else either in part or in full and it is the original work, for the award of any other degree or diploma in this institute or any other institute or university.

**Prof. Shivangi Mehta**

**Assistant Professor**

**Devang Patel Institute of Advanced Technology And Research FTE, CHARUSAT-Changa.**



### CERTIFICATE

This is to certify that the report entitled “Phishing Site Detection using Machine Learning” is a bonafied work carried out by **Mr. Vaibhav Kanapariya (19DCS051), Mr Shrey Gajjar (19DCS034)** under the guidance and supervision of Asst. Professor **Shivangi Mehta** for the subject **Software Group Project-IV** (CSE) of 6th Semester of Bachelor of Technology in **DEPSTAR** at Faculty of Technology & Engineering – CHARUSAT, Gujarat.

To the best of my knowledge and belief, this work embodies the work of candidate himself, has duly been completed, and fulfills the requirement of the ordinance relating to the B.Tech. Degree of the University and is up to the standard in respect of content, presentation and language for being referred to the examiner.

|  |
| --- |
| Prof **Shivangi Mehta**(CSE)  DEPSTAR, Changa, Gujarat. |
| Dr. Amit Ganatra Principal, DEPSTAR Dean, FTE  CHARUSAT, Changa, Gujarat. |
|  |

**Devang Patel Institute of Advance Technology And Research At: Changa, Ta. Petlad, Dist.**

**Anand, PIN: 388 421. Gujarat**

### ABSTRACT

Now a days there are lot of attacks are going on to steal someone’s information or data. Hackers have advanced to the point where they can use their knowledge to break into someone else's system and steal information. Criminals have discovered a way to steal personal information without having to meet them and with the least risk of being caught. It is called Phishing. Phishing is a cybercrime that targets emails, phone calls, text messages, personally identifiable information, banking information, credit card information, and passwords. Phishing is essentially an online identity theft scheme. The phisher uses Social Engineering to steal the victim's personal information and account information. This paper gives awareness about phishing attacks and anti-phishing tools,types of phishing attacks through which the attacks are performed, detection and prevention towards it.

### ACKNOWLEDGEMENT

We thanks Mr. **Shivangi Mehta** (Asst. Professor) who have been the great inspiration and who have provided sufficient background knowledge and understanding of this subject.

Our humble prostration goes to our H.O.D. Dr. **Parth Goel** , for the support during the whole session of study and development. It is because of them, that we were prompted to do hard work, adopting new technologies which have aided us to complete this project successfully.

we would also like to thank our mentor Prof. **Shivangi Mehta** for his guidelines throughout the development phase of the project.

They altogether provided me favourable environment, and without them it would not have been possible to achieve my goal.

Thanks,

Vaibhav Kanapariya Shrey Gajjar

**INDEX**

Declaration By The Candidate 2

Certificate 3

Abstract 4

Acknowledgement 5

1. [INTRODUCTION](#_bookmark0) 9
2. [System Requirements](#_bookmark1) 10
3. [System Flowchart and project working](#_bookmark3) 12
4. [Screenshots of the project output](#_bookmark4) 15
5. [Limitation Of The Project](#_bookmark14) 19
6. [Project Outcomes](#_bookmark15) 18
7. [Future Enhancement](#_bookmark16) 20
8. [Conclusion](#_bookmark17) 21
9. [References](#_bookmark18) 22

# CHAPTER 1: INTRODUCTION

### PROJECT DEFINITION

Phishing is an act of attempting a victim for fraudulently acquires sensitive information such as username, password, bank details, credit card details or some valuable information. Phishing attacks are designed to trick recipients into disclosing sensitive information such as bank account numbers, passwords, and credit card numbers. For example, despite knowing little to nothing about the recipient, a phisher may misrepresent himself as a large banking corporation or a popular on-line auction site will have a reasonable yield. Phishing emails may contain links to websites that are infected with malware. Phishing is mostly utilised in email hacking; in email phishing, the hacker sends a link to the user with, say, bank account information via email. or any other personal information, the user is now directed to that site. He fills in all of the information on that link, after which the hacker obtains all of the information. This is how phishing is done.

Some commercial anti-spam and anti-phishing products block email from "blacklisted" sites that they claim deliver spam and phishing emails, but allow email from "whitelisted" sites that they claim don't send it. Phishing is explained step-by-step:

1. Attacker sends an email to victim.
2. Victim clicks on the email and goes to phishing website.
3. Attacker collects victim’s credentials.
4. Attacker uses victim’s credentials to access a website.

Phishing begins with an email or other form of communication intended to aid in the victim's attack. The communication is made to appear as if it was sent by a trusted sender. If the victim is duped, the victim will provide personal information to a spam website. Malware may potentially be downloaded onto the target's PC.

### OBJECTIVE OF PROJECT.

Nowadays privacy and security is very important to everyone. Because every one private details is interconnected to each other if third person somehow access the one detail he might be get all the detail of user.

Due to this, we have created the model which help user weather the given site or application is phishing site or not.

In our model user have to give website URL which user want to check and prints whether given URL is phishing site or not.

### TOOLS AND TECHNOLOGY USED

* + - We have used Google Colaboratory notebook Software
    - Fast API
    - Python libraries
    - Kaggle datasets

# CHAPTER 2:SYSTEM REQUIREMENTS

## Hardware And Software Requirements:

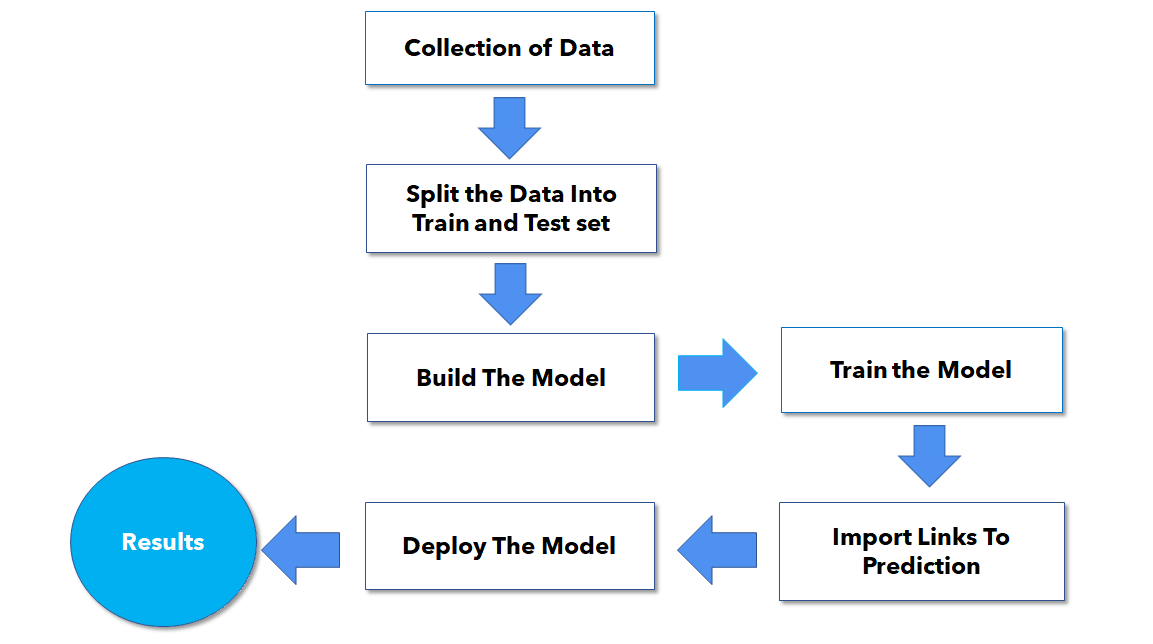
* + - Operating system: Mac OS, Windows
    - Internet connection with good speed

# CHAPTER 3 :SYSTEM FLOW CHART AND WORKING

### WORKING

* Phishing is the method by which fraudulent information is sent to victims which appears to arrive from the legitimate sources. Mostly, the attackers chosen mode of delivering the infection is through email because it is more effective. The purpose is to rob the user of their sensitive data like credit card and login information or to install malware on the victim’s machine.
* Phishing starts with a fraudulent email or other communication that is designed to lure a victim. The message is made to look as though it comes from a trusted sender. If it fools the victim, he or she is coaxed into providing confidential information, often on a scam website. Sometimes malware is also downloaded onto the target’s computer.
* Here are the basic steps to check whether the site is phishing site or not:
* Enter the site URL which you want to check whether the site is phishing site or not.
* The machine learning model checks the url and identify it that whether it is good or bad.
* The model checks the sites and finds the keywords which are present in the url and checks it from the database.
* From this you can simply check the site and secure yourself from attack

**Flowchart:**



# CHAPTER 4: SCREENSHOTS OF PROJECT OUTPUT

### SCREENSHOTS:

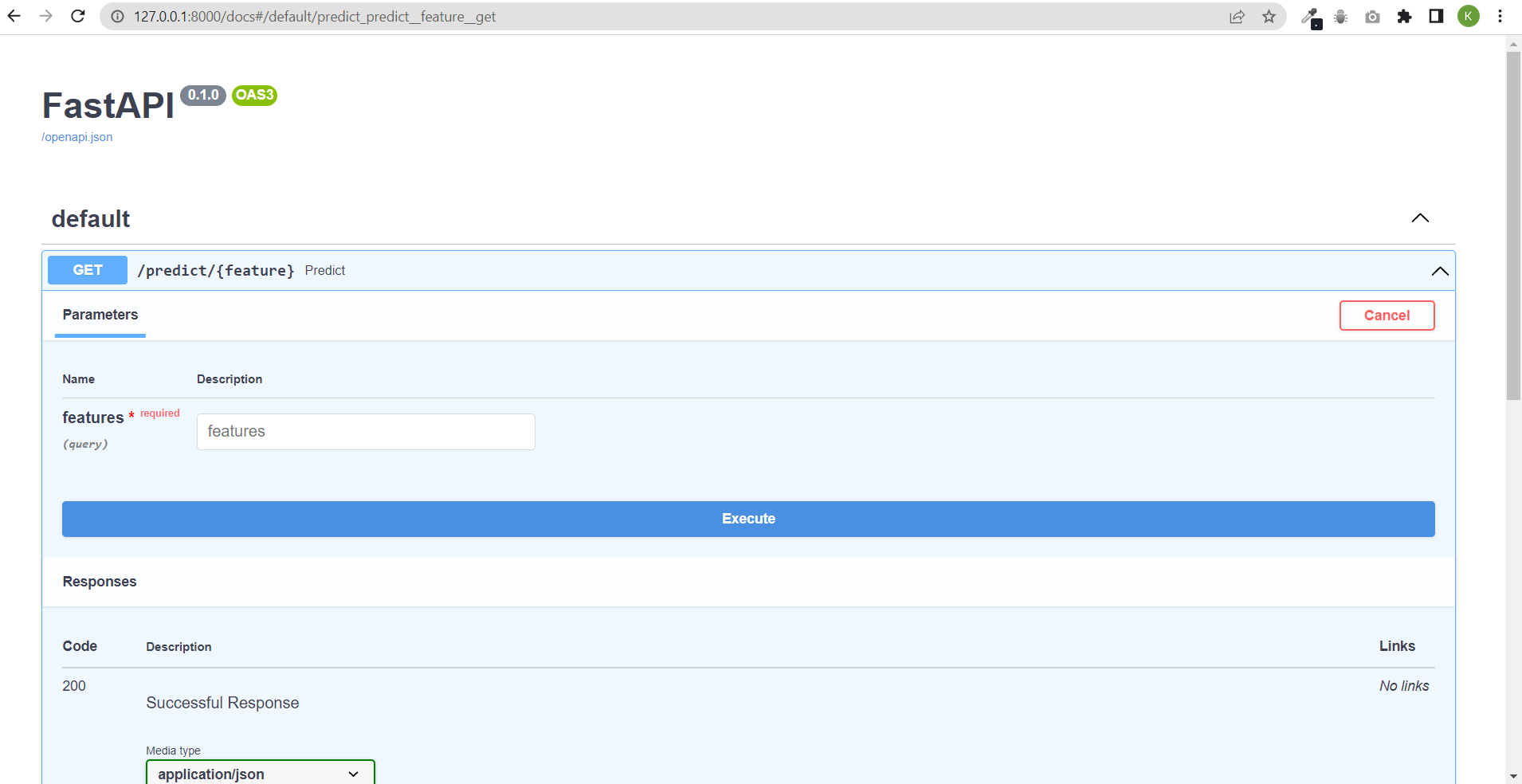
### Google Colaboratory notebook :

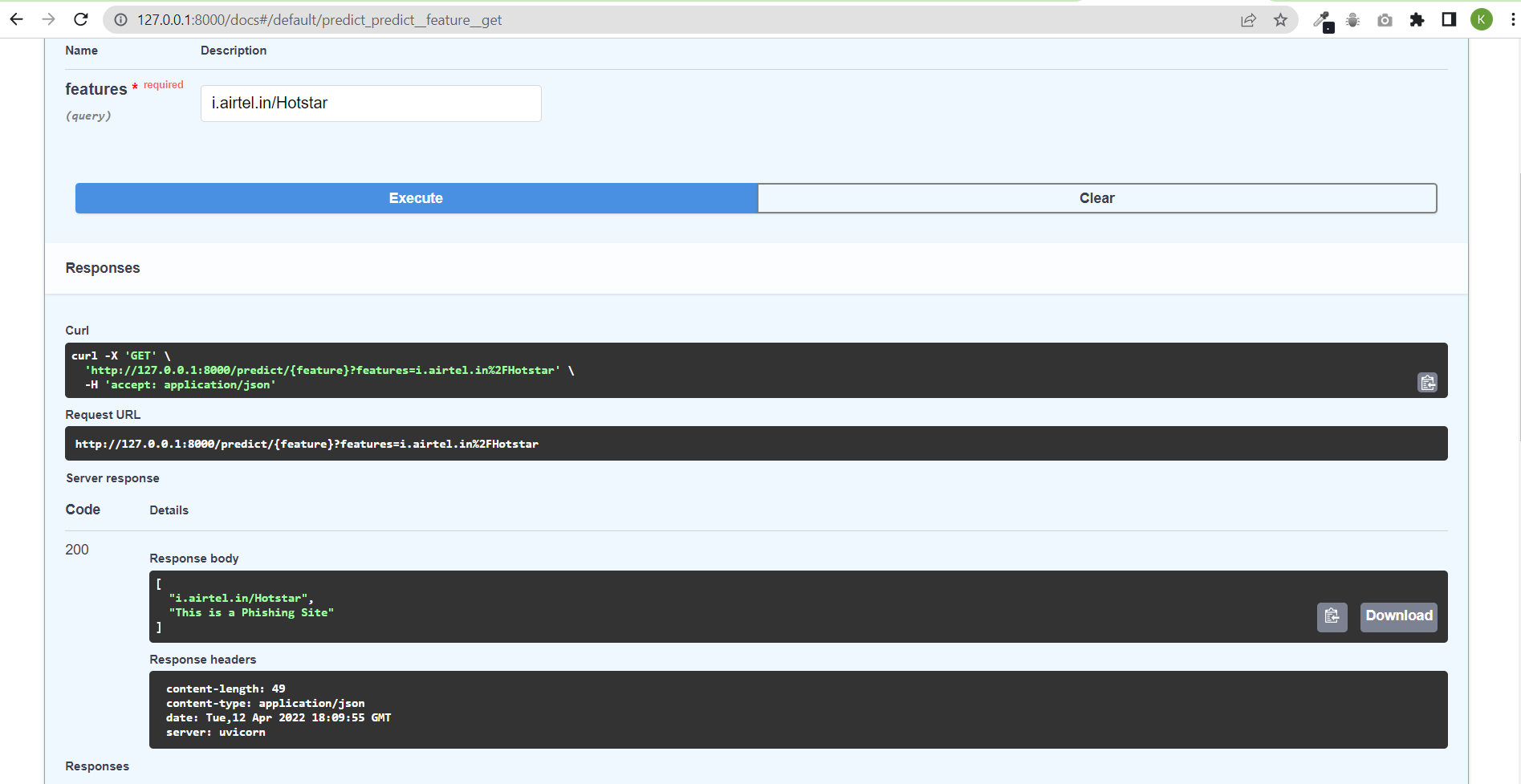
### A screenshot of a computer Description automatically generated

Graphical user interface, text, application

Description automatically generated

**FastAPI:**





**CHAPTER 5: LIMITATION OF THE PROJECT**

* It can’t be run in background as chrome extension.
* It might not identify some URL.
* Not good looking UI.
* Not hosted website.

## CHAPTER 6: PROJECT OUTCOMES

* + Easley identify whether the site is phishing site or not.
  + Prevent user from losing their crucial information/data.
  + Give user warning not to go further.

## CHAPTER: 7 : FUTURE ENHANCEMENT

* Make a chrome extension which runs in background.
* Making mobile application which runs in background.
* Make model more accurate to detect other files also.
* Also train the model to prevent the other cyber attacks also.

## CHAPTER: 8 : CONCLUSION

Phishing is a method that uses fraudulent URLs and emails to acquire sensitive information about the victim. It is one of the most hazardous cyber-attacks that take place in corporations, personal gadgets, and other places. It is frequently difficult to distinguish between legitimate and phishing emails. This assault may be avoided using a variety of approaches. Anti-phishing solutions and platforms that are regularly updated can be quite effective. This paper gives an understanding of phishing, the process of the assault, the numerous forms it may take, and viable strategies to overcome them.

## CHAPTER: 9 : REFERENCES

* <https://www.frontiersin.org/articles/10.3389/fcomp.2021.563060/full>
* <https://www.researchgate.net/publication/226420039_Detection_of_Phishing_Attacks_A_Machine_Learning_Approach>
* <https://www.youtube.com/watch?v=biKDqy-_J1o>
* https://www.kaggle.com/