

VISVESVARAYA TECHNOLOGICAL UNIVERSITY

“JNANA SANGAMA”, BELAGAVI-590018



Application Development Using Python Mini Project Report On

“PHONE NUMBER & EMAIL EXTRACTOR”

SUBMITTED IN PARTIAL FULFILMENT FOR 5TH SEMESTER.

BACHELOR OF ENGINEERING IN COMPUTER SCIENCE AND ENGINEERING

SUBMITTED BY

**VARSHITHA Y S(1JB20CS134)
SUSHMA H J(1JB20CS128)**

UNDER THE GUIDANCE OF

Mr. Arun Kumar

Assistant Professor

Dept. of CSE,

SJBIT, Bengaluru-60



DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING

SJB INSTITUTE OF TECHNOLOGY

**#No.67, BGS HEALTH & EDUCATION CITY, DR. VISHNUVARDHAN ROAD, KENGARI,
BENGALURU-560060, KARNATAKA, INDIA.**

2022-23

|| Jai Sri Gurudev ||
Sri Adichunchanagiri Shikshana Trust ®

SJB INSTITUTE OF TECHNOLOGY

No.67, BGS Health & Education City, Dr. Vishnuvardhan Rd, Kengeri, Bengaluru, Karnataka 560060

Department of Computer Science and Engineering



CERTIFICATE

Certified that the Application Development Using Python Mini Project entitled “**PHONE NUMBER AND EMAIL EXTRACTOR**” carried out by **Varshitha Y S, Sushma H J** bearing USN **1JB20CS134 and 1JB20CS128** are bonafide students of **SJB Institute of Technology** in partial fulfilment for 5th Semester of **BACHELOR OF ENGINEERING** in **COMPUTER SCIENCE AND ENGINEERING** of the **Visvesvaraya Technological University, Belagavi** during the academic year **2022-23**. It is certified that all corrections/suggestions indicated for Internal Assessment have been incorporated in the Report deposited in the Departmental library. The mini project report has been approved as it satisfies the academic requirements in respect of mini project prescribed for the said Degree.

Signature of Guide

Mr. Arun Kumar

Assistant professor

Dept. of CSE, SJBIT

Signature of HOD

Dr. Krishna A N

Professor & Head

Dept. of CSE, SJBIT



ACKNOWLEDGEMENT

We would like to express our profound grateful to His Divine Soul **Jagadguru Padmabhushan Sri Sri Sri Dr. Balagangadharanatha Mahaswamiji** and His Holiness **Jagadguru Sri Sri Sri Dr. Nirmalanandanatha Mahaswamiji** for providing us an opportunity to complete our academics in this esteemed institution.

We would also like to express our profound thanks to **Revered Sri Sri Dr. Prakashnath Swamiji**, Managing Director, SJB Institute of Technology, for his continuous support in providing amenities to carry out this Mini Project in this admired institution.

We express our gratitude to **Dr. K. V. Mahendra Prashanth**, Principal, SJB Institute of Technology, for providing us an excellent facility and academic ambience; which have helped us in satisfactory completion of Mini Project work.

We extend our sincere thanks to **Dr. Krishan A N**, Head of the Department, Computer Science and Engineering for providing us an invaluable support throughout the period of our Mini project work.

We wish to express our heartfelt gratitude to our guide **Mr. Arun Kumar** Assistant Professor Dept. Of CSE SJBIT for her valuable guidance, suggestions, and cheerful encouragement during the entire period of our Mini Project work.

Finally, we take this opportunity to extend our earnest gratitude and respect to our parents, Teaching & Non-teaching staffs of the department, the library staff and all our friends, who have directly or indirectly supported us during the period of our Mini Project work.

Regards,

Varshitha Y S [1JB20CS134]

Sushma H J [1JB20CS128]

INTRODUCTION

If we have the boring task of finding every phone number and email address in a long web page or document. If you manually scroll through the page, you might end up searching for a long time. But if you had a program that could search the text in your clipboard for phone numbers and email addresses, you could simply press ctrl-A to select all the text, press ctrl-C to copy it to the clipboard, and then run this project. It could replace the text on the clipboard with just the phone numbers and email addresses it finds.

A Regular Expression (RegEx) is a special sequence of characters that uses a search pattern to find a string or set of strings. It can detect the presence or absence of a text by matching it with a particular pattern, and also can split a pattern into one or more sub-patterns. Python provides a `re` module that supports the use of regex in Python. Its primary function is to offer a search, where it takes a regular expression and a string. Here, it either returns the first match or else none.

`Pyperclip` is a cross-platform Python module for copy and paste clipboard functions. This module was created to enable cross-platform copy-pasting in Python which was earlier absent. The `pyperclip` module has `copy()` and `paste()` functions that can send text to and receive text from your computer's clipboard.

IMPLEMENTATION

Requirements

- Editor (VS code)
- Any editor supporting Python Compilation
- Python latest version
- Hardware: Processor intel
- Hardware: 8GB RAM

Working of Project

- Get the text off the clipboard.
- Find all phone numbers and email addresses in the text.
- Paste them onto the clipboard.

The code will do the following:

- Use the pyperclip module to copy and paste strings.
- Create two regexes, one for matching phone numbers and the other for matching email addresses.
- Find all matches, not just the first match, of both regexes.
- Neatly format the matched strings into a single string to paste.
- Display some kind of message if no matches were found in the text

SOURCE CODE

```
#to extract Indian phone number and email
```

```
import pyperclip, re
```

```
IndianNumber = re.compile(r"(
(?:\d{0,2})91(\s*[-]\s*)?|[0]?)?[789]\d{9}
)",re.VERBOSE)
```

```
emailRegex = re.compile(r"(
[a-zA-Z0-9._%+-]+      # Username
@                      # @Symbol
[a-zA-Z0-9.-]+        # Domain Name
(\.[a-zA-Z]{2,4})     # dot-something
)", re.VERBOSE)
```

```
# Find Matches in Clipboard Text
```

```
text = str(pyperclip.paste())
```

```
email_groups = emailRegex.findall(text)
```

```
Indian_Contacts = IndianNumber.findall(text)
```

```
matched = []
```

```
for group in Indian_Contacts:
```

```
    phoneNum = group[0]
```

```
    matched.append(phoneNum)
```

```
for group in email_groups:
```

```
    matched.append(group[0])
```

```
if len(matched) > 0:
```

```
    pyperclip.copy("\n".join(matched))
```

```
    print('Copied to clipboard!\n')
```

```
    print("\n".join(matched))
```

```
else:
```

```
    print('No Phone Numbers or Emails found in your website')
```

OUTPUT

- Open any webpage ex: <https://www.sjbit.edu.in/sjbit-contacts.html>.
- press ctrl-A to select all the text, press ctrl-C to copy it to the clipboard.
- Run the project.
- It displays below output in our terminal for respective webpage address.

Copied to clipboard!

+91-9900638899
+91-7259855234
+91-7676554550
+91-9449213292
9986502012
+91-7795335147
9880762882
+91-8711065562
9035649034
9916065415
9916101233
9916101204
9739099708
9739099709
9739099706
9739099707
9632354173
9449318776
9739788858
9739801965
9449379213
9886811828
+91-7204081185
+91-9902310181
principal@sjbit.edu.in
ecchod@sjbit.edu.in
mechhod@sjbit.edu.in
civhod@sjbit.edu.in
csehod@sjbit.edu.in
isehod@sjbit.edu.in
eeehod@sjbit.edu.in
kvmprashanth@sjbit.edu.in
mathshod@sjbit.edu.in
phyhod@sjbit.edu.in
chehod@sjbit.edu.in
mbahod@sjbit.edu.in
sports@sjbit.edu.in
librarian@sjbit.edu.in
principal@sjbit.edu.in