



Shri Yashwantrao Bhonsale Education Society's
YASHWANTRAO BHONSALE INSTITUTE OF TECHNOLOGY

(DTE CODE : 3470) (MSBTE Code : 1742)

Approved by AICTE, DTE & Affiliated to Mumbai University & MSBTE Mumbai
(NBA Accredited ME, CE, EE Diploma Programs)

Practical No 17

Aim

Create a program that reads a text file containing various data (e.g., names, emails, phone numbers). Use regular expressions to extract specific types of data.

Apparatus / Software Required

- Python Interpreter (Python 3.14.2)

Theory

◆ Regular Expressions (Regex) are used in Python for searching and extracting specific patterns from text data. The built-in module `re` module provides functions to perform pattern matching.

In this practical, a text file containing mixed data such as:

- Names
- Email IDs
- Phone Numbers

is read, and Regex is used to extract specific information like emails and phone numbers.

This is useful in real-world applications like:

- Data mining
- Log analysis
- Form data processing

◆ Important Methods Used

★ 1. `re.findall()` (Most Important)

- Returns all matches of a pattern in the text
- Output is in the form of a list

`re.findall(pattern, text)`



Shri Yashwantrao Bhonsale Education Society's
YASHWANTRAO BHONSALE INSTITUTE OF TECHNOLOGY

(DTE CODE : 3470) (MSBTE Code : 1742)

Approved by AICTE, DTE & Affiliated to Mumbai University & MSBTE Mumbai
(NBA Accredited ME, CE, EE Diploma Programs)

★ 2. **re.search()**

- Finds the first occurrence of a pattern
- Returns match object

◆ 3. **re.match()** (Less Important Here)

- Matches pattern only at the beginning

◆ 4. **open()** (File Handling Function)

- Used to read data from a file

```
file = open("filename.txt", "r")
```

✉ Regex Patterns Used

Email Pattern:

```
[a-zA-Z0-9._%+-]+@[a-zA-Z0-9.-]+\.[a-zA-Z]{2,}
```

Phone Number Pattern:

```
(\+91)?[6-9]\d{9}
```

📄 Algorithm

1. Start
2. Import **re** module
3. Open the text file in read mode
4. Read file content
5. Define regex pattern for email
6. Define regex pattern for phone number
7. Use **re.findall()** to extract emails
8. Use **re.findall()** to extract phone numbers
9. Display extracted data
10. Stop

