```
import re
```

```
testdate = '''Possible dates for the first assessment exams are 1-08-2018, 03/8/20
Possible dates for the second assessment exams are October 22 , 2018 ,  October 24
1 Nov,2018, 03 Nov ,2018 32 Nov,2018 Possible dates for semester exams are 2018/11,


print(re.findall('\\b\d{1,2}/\d{1,2}/\d{2,4}\\b', testdate)) #dd/mm/yy or #dd/mm/y

print(re.findall('\\b\d{1,2}\.\d{1,2}\.\d{2,4}\\b', testdate)) #dd.mm.yyyy or #dd.

print(re.findall('\\b\d{4}-\d{1,2}-\d{1,2}\\b', testdate)) #yyyy-mm-dd

print(re.findall('(?:January|February|October) \d{1,2} *\, *\d{4}\\b', testdate)) #

print(re.findall('\d{1,2} (?:Jan|Feb|Mar|Apr|May|June|July|Aug|Sept|Oct|Nov|Dec) *'
```

```
['03/8/2018']
['07.08.18']
['2018-11-9', '2018-13-04']
['October 22 , 2018', 'October 24, 2018', 'October 29,  2018']
['1 Nov,2018', '03 Nov ,2018', '32 Nov,2018']
```

```
testtime = '''Possible starting  times are
5.15AM  or 05:15 AM  or 5.15am or 5.15 am
5.15 PM or 5:15PM  or 5.15 PM or 5.15pm
13:4 or 13:04:22 or 13:24:22.04'''

#4.30 AM, 4:30 am, 5.15 pm, 13:24

print(re.findall('\\b\d{1,2}\.\d{1,2} AM',testtime)) #H.MM AM

print(re.findall('\\b\d{1,2}\:\d{1,2} am',testtime)) #H:MM am

print(re.findall('\\b\d{1,2}\.\d{1,2} PM',testtime)) #H.MM PM

print(re.findall('\\b\d{2}\:\d{2}',testtime)) #HH:MM
```

```
[]
[]
['5.15 PM', '5.15 PM']
['05:15', '13:04', '13:24']
```

```
testmob = '''Possible  mobile number formats are 7382293451
91738293451 or 91 7382293451 +917382293451 or +91 7382293451 or
+  91 7382293451 (+91)738229345 or ( +91 )7382293451 or (+91) 7382293451'''

#7382293451 +917382293451 (+91) 738229345

print(re.findall('[^\+)\d] \d{10}',testmob))

print(re.findall('\+\d{2}\d{10}',testmob))

print(re.findall('\(\+91\) \d{10}', testmob))
```

```
['e 7382293451']
['+917382293451']
['(+91) 7382293451']
```

```
testll = '''Landline number formats are 22469 04322469 or 043 22469 (043)22469 or
```

```
#22469  04322469  (043)22469

print(re.findall('[^)\d] \d{5}',testll))

print(re.findall('043\d{5}',testll))

print(re.findall('\(043\)\d{5}',testll))
```

```
['e 22469']
['04322469']
['(043)22469']
```

```
testcur = '''$500.12 , 500.12dollars, 500.1 dollars
$ 1.24 , 1.25dollars,  $1,000.40 , $10,000.14'''

#10 dollars    $ 10

print(re.findall('\d+\.\d* dollars',testcur))

print(re.findall('\$ \d+\.\d*',testcur))
```

```
['500.1 dollars']
['$ 1.24']
```

```
#1046 → 1,046
#100049 → 1,00,49 [Indian]
#1000496 → 1,000,496 [US]

testcurcon = '''1046'''
print(re.sub('([\d])(\d{3})','\\1,\\2',testcurcon))

testcurcon2 = '''100049'''
print(re.sub('([\d])(\d{2})(\d{2})','\\1,\\2,\\3',testcurcon2))

testcurcon3 = '''1000496'''
print(re.sub('([\d])(\d{3})(\d{3})','\\1,\\2,\\3',testcurcon3))
```

```
1,046
1,00,049
1,000,496
```

```
testmails = '''abcd.yz [at] iiith [dot] ac[dot]
in  can be written like  abcd.yz [AT] iiith [DOT] ac [DOT] in  or
like  abcd.yz(AT) iiith (DOT) ac (DOT) in , abcd.yz(at ) iiith ( dot) ac. In,
abcd[at]nitt.edu  or  abcd AT nitt DOT edu  etc'''

#abcd.yz [at] iiith [dot] ac [dot] in
#abcd [at] nitt.edu
#abcd AT nitt DOT edu
#abcd.yz(AT) iiith (DOT) ac (DOT) in

print(re.findall('\w+\.\w+ \[at\] iiith \[dot\] ac \[dot\] in',testmails))

print(re.findall('\w+ \[at\] nitt.edu',testmails))

print(re.findall('\w+ AT nitt DOT edu', testmails))

print(re.findall('\w+\.\w+\(AT\) iiith \(DOT\) ac \(DOT\) in', testmails))
```

```
[]
[]
['abcd AT nitt DOT edu']
['abcd.yz(AT) iiith (DOT) ac (DOT) in']
```

```
testurls = '''https://www.google.co.in www.cricbuzz.in http://www.abcd.me
This course is very good E:\nlp\deep learning\ Andrew ng\ course1\introduction.mp4
```

```
#www.website.extension
#http[s]://www.website.extenstion

print(re.findall('[^/]www\.\w+\.\w+', testurls))

print(re.findall('https?\://www\.\w+\.\w+\.?\w*',testurls))
```

```
[' www.cricbuzz.in']
['https://www.google.co.in', 'http://www.abcd.me']
```

```
testnames = '''Rama Viswanath called by her friends Rama V. or V.Rama Subash Chand
called as  S. C. Bose  or sometimes  Subash C. Bose, Mr. Subash Chandra Bose, Mr. S
Sri  A. Sumanth,  Dr. S. Priya  , Dr.B.Kavitha, Prof. John Xavier are members of t


print(re.findall('[a-zA-z]{4,} [a-zA-Z]{3,}' ,testnames)) #Rama Viswanath  #false

print(re.findall('\S{4,} \S\.' ,testnames)) #Rama V.

print(re.findall('[a-zA-Z]{1}\.[a-zA-Z]{4,}' ,testnames)) #V.Rama

print(re.findall('[a-zA-z]{4,} [a-zA-Z]{3,} [a-zA-Z]{4,}', testnames)) #Subash Cha

print(re.findall('\\b[a-zA-Z]{1}\. [a-zA-Z]{1}\. [a-zA-Z]{4,}' ,testnames)) #S. C.

print(re.findall('\\b[a-zA-Z]{3,} [a-zA-Z]{1}\. [a-zA-Z]{4,}' ,testnames)) #Subash

print(re.findall('Mr\. [a-zA-z]{4,} [a-zA-Z]{3,} [a-zA-Z]{4,}', testnames)) #Mr. S

print(re.findall('Mr\. [a-zA-Z]{1}\. [a-zA-Z]{1}\. [a-zA-Z]{4,}' ,testnames)) #Mr.
```

```
['Rama Viswanath', 'friends Rama', 'Rama Subash', 'Chandra Bose', 'Subash Ch
['Rama V.', 'Subash C.']
['V.Rama', 'B.Kavitha']
['Rama Viswanath called', 'Rama Subash Chandra', 'Subash Chandra Bose', 'Xav
['S. C. Bose', 'S. C. Bose']
['Subash C. Bose']
['Mr. Subash Chandra Bose']
['Mr. S. C. Bose']
```

```
testabr = "U.S.A. and I.T. and U.S.A."
print(re.sub('[A-Z]\.[A-Z]\.[A-Z].','United States of America', testabr)) #3 lette
print(re.sub('[^A-Z\.][A-Z]\.[A-Z]\.[^A-Z]','Information Technology', testabr)) #2
```

```
United States of America and I.T. and United States of America
U.S.A. andInformation Technologyand U.S.A.
```