In [1]:
```
import re

def filter_regional_language(text, language_pattern):
    filtered_text = re.sub(language_pattern, '', text)
    return filtered_text

hindi_pattern = re.compile(r'[\u0900-\u097F]+')
input_sentence = input("Enter a sentence: ")
print("Input Sentence:", input_sentence)
filtered_sentence = filter_regional_language(input_sentence, hindi_pattern)
print("Filtered Sentence:", filtered_sentence)
```

```
Enter a sentence: Hello.  कैसे हो?
Input Sentence: Hello.  कैसे हो?
Filtered Sentence: Hello.   ?
```

In [2]:
```
#Stop word Filtraton

import re

def filter_stop_words(sentence) :
    stop_words = set([
        "i", "me", "my", "myself", "we", "our", "ours","you", "your",
        "yours", "he", "him", "his", "himself", "she", "her", "hers", "it",
        "its", "itself", "they", "them", "their", "theirs", "what",
        "which", "who", "whom", "this", "that", "these", "those", "am",
        "is", "are", "was", "were", "be", "been", "being", "have", "has",
        "had", "having", "do","through", "during", "before", "after", "above",
        "below", "to", "from", "up", "down", "in", "out", "on", "off", "over", "und
        "again", "further", "then", "once", "here", "there", "when", "where",
        "why", "how", "all", "any", "both", "each", "few", "more", "most",
        "other", "some", "such", "no", "nor", "not", "only", "own", "same",
        "so", "than", "too", "very", "s", "t", "can", "will", "just", "don",
        "should", "now"
    ])

    word_pattern = re.compile(r'\b\w+\b')
    filtered_sentence = word_pattern.sub(lambda match: match.group() if match.group
    return filtered_sentence

input_sentence = input("Enter a sentence: ")
filtered_sentence = filter_stop_words(input_sentence)
print("Filtered Sentence:", filtered_sentence)
```

```
Enter a sentence: This is the first experiment of NLP. It is a subject for AI/DS d
epartment students in sem 7. There are some other subjects such as DL, BDA and a f
ew more. This above cell executes stop word filtration.
Filtered Sentence:   the first experiment of NLP.   a subject for AI/DS department
students  sem 7.     subjects  as DL, BDA and a  .   cell executes stop word filtr
ation.
```

In [3]:
```
# Punctuation FIltration

import re

def filter_punctuation(sentence) :
    filtered_sentence = re.sub(r'[^\w\s]', '', sentence)
    return filtered_sentence

input_sentence = input("Enter a sentence: ")
filtered_sentence = filter_punctuation(input_sentence)
print("Filtered Sentence:", filtered_sentence)
```

```
Enter a sentence: Hello! Welcome world, how are you? Have a good day.
Filtered Sentence: Hello Welcome world how are you Have a good day
```

In [4]:
```python
#Email verification

import re

def is_valid_email(email) :
    pattern = r'^[a-zA-Z0-9._%+-]+@[a-zA-Z0-9.-]+\.[a-zA-Z]{2,}$'
    return re.match(pattern, email)

input_email = input("Enter an email address: ")

if is_valid_email(input_email) :
   print("Email is valid.")
else :
   print("Email is not valid.")
```

```
Enter an email address: subrato.3.gmail.com
Email is not valid.
```

In [5]:
```python
# Phone no validation

import re

def is_valid_phone_number(phone_number) :
    # Regular expression for basic phone number format validation
    pattern = r'^[2-9]\d{2}-\d{3}-\d{4}$'
    return re.match(pattern, phone_number)

input_phone_number = input("Enter a phone number : ")

if is_valid_phone_number(input_phone_number) :
  print("Phone number is valid.")
else :
  print("Phone number is not valid.")
```

```
Enter a phone number : 145879632a
Phone number is not valid.
```

In [6]:
```python
# Name validation

import re

def is_valid_name(name) :
    # Regular expression for name validation (letters and spaces only)
    pattern = r'^[a-zA-Z\s]+$'
    return re.match(pattern, name)

input_name = input("Enter a name : ")

if is_valid_name(input_name) :
    print("Name is valid.")
else :
    print("Name is not valid.")
```

```
Enter a name : Subrato
Name is valid.
```

In [7]:
```python
input = input("Enter a sentence : ")
```

```
Enter a sentence : hello how are you
```

In [8]:
```python
# Tokenization
```

```
list = []
for item in input.split(" ") :
  list.append(item)
print(list)
```

```
['hello', 'how', 'are', 'you']
```