

NATIONAL UNIVERSITY OF MODERN LANGUAGES
ISLAMABAD



Natural Language Processing

Submitted to:

Ms. Qurat –ul- ain Safder

Submitted By:

Nida Shafiq(BSAI-139)

Aleeha Zahid(BSAI-136)

Azalfa Awan (BSAI-135)

Nida Niazi(BSAI-147)

Sahab Mushtaq(BSAI-155)

Dania Safreen(BSAI-061)

Submission Date: March 07, 2025

1. Create a regular expression that matches all words containing exactly three vowels.

Here is the regular expression that achieves this:

```
\b[^aeiouAEIOU]*[aeiouAEIOU][^aeiouAEIOU]*[aeiouAEIOU][^aeiouAEIOU]*[aeiouAEIOU]\b
```

Explanation:

- ``\b`` → Word boundary (ensures we're matching whole words).
- ``[^aeiouAEIOU]*`` → Matches any number of consonants before or between vowels.
- ``[aeiouAEIOU]`` → Matches a vowel.
- This pattern repeats three times to ensure exactly three vowels.
- ``\b`` → Ends at a word boundary.

This ensures that the word contains exactly three vowels, with any number of consonants in between.

2. What is the role of anchors (^ and \$) in regular expressions?

Answer: Anchors are used to match positions within a string rather than actual characters.

- The caret (^) matches the start of a string. For example, `^Hello` matches "Hello" only if it appears at the beginning of a line.
 - The dollar sign (\$) matches the end of a string. For example, `world$` matches "world" only if it appears at the end of a line.
-

3. How does the question mark (?) operator work in regular expressions?

Answer: The question mark (?) makes the preceding character or group optional, meaning it can appear **zero or one** time.

Example:

- `colou?r` matches both "color" and "colour" because the u is optional.
-

4. What is the difference between the Kleene star (*) and the question mark (?) in regex?

Answer:

- The **Kleene star (*)** matches **zero or more** occurrences of the preceding element. Example: `a*` matches "", "a", "aa", "aaa", etc.

- The **question mark (?)** matches **zero or one** occurrence. Example: `a?` matches "" or "a" but not "aa".
-

5. What is the purpose of disjunction (|) in regular expressions?

Answer: The disjunction operator (|) acts like an OR operator, allowing a match with **either** of the specified patterns.

Example:

- `cat|dog` matches either "cat" or "dog".
 - `(apple|orange|banana)` matches "apple", "orange", or "banana".
-

6. How does grouping () affect precedence in regular expressions?

Answer: Parentheses () are used to **group expressions** and **change precedence** in regex.

- Without grouping: `a|bc` matches "a" OR "bc".
- With grouping: `(a|b)c` matches "ac" OR "bc".
They also allow for capturing and referencing matched subpatterns.