

## Labeled Dataset( $C$ )

Precprocessed-data( $d$ )	Outcome( $l$ )
"আমি ভালো আছি"	Normal
"মন খারাপ লাগছে আমার"	Sick
"আজকে অনেক ভালো লাগছে"	Normal
"আমার মাথা ব্যথা অনেক"	Sick

## Keyword-Lists

- **Positive-keywords-list (Label-Normal): ( $V_{pos}$ )**  
["আমি", "ভালো", "আছি", "আজকে", "অনেক", "লাগছে"]
  - **Negative-keywords-list (Label-Sick): ( $V_{neg}$ )**  
["মন", "খারাপ", "লাগছে", "আমার", "মাথা", "ব্যথা", "অনেক"]
- 

## Example:

### Input:

"আজকে আমার মন খারাপ"

### Step-by-step-process:

1. **Splitting words:**  
["আজকে", "আমার", "মন", "খারাপ"]
  2. **Counting positive frequency:**  
 $V_{pos} = ["আমি", "ভালো", "আছি", "আজকে", "অনেক", "লাগছে"]$ 
    - Found "আজকে". Positive frequency = 1
  3. **Counting negative frequency:**  
 $V_{neg} = ["মন", "খারাপ", "লাগছে", "আমার", "মাথা", "ব্যথা", "অনেক"]$ 
    - Found "আমার", "মন", "খারাপ ". Negative frequency = 3
- 

## Feature Vector:

[1, Positive\_frequency, Negative\_frequency] = [1, 1, 3]