LetsGrowMore (LGM) - virtual internship program

Name: Shiva Dagdu Mehenge

Data Science Intern

Task:- Next Word Prediction:

```python
import numpy as np

lexicon = {}

def update_lexicon(current : str, next_word : str) -> None:
    # Add the input word to the lexicon if it in there yet.
    if current not in lexicon:
        lexicon.update({current: {next_word: 1} })
        return

    # Recieve te probabilties of the input word.
    options = lexicon[current]

    # Check if the output word is in the propability list.
    if next_word not in options:
        options.update({next_word : 1})
    else:
        options.update({next_word : options[next_word] + 1})

    # Update the lexicon
    lexicon[current] = options
```

Load Dataset

```
In [3]:   1  with open('C:\\Users\\Shivv\\dataset.txt', 'r', encoding ='ISO-8859-1') as dataset:
          2      for line in dataset:
          3          words = line.strip().split(' ')
          4          for i in range(len(words) - 1):
          5              update_lexicon(words[i], words[i+1])
```

Train Model | Predicting the next Word

```
In [8]:   1  for word, transition in lexicon.items():
          2      transition = dict((key, value / sum(transition.values())) for key, value in transition.items())
          3      lexicon[word] = transition
          4
          5  line = input('> ')
          6  word = line.strip().split(' ')[-1]
          7  if word not in lexicon:
          8      print('Word not found')
          9  else:
         10      options = lexicon[word]
         11      predicted = np.random.choice(list(options.keys()), p=list(options.values()))
         12      print(line + ' ' + predicted)
```

```
> i love of my
i love of my confidence.
```

```
In [ ]:   1
```