Odyssey Lab - Task

What is a Markov Chain?

A Markov Chain is a mathematical system that undergoes transitions from one state to another, within a finite or countable number of possible states. It is a random process that has the key property of Markov property, meaning that the future state depends only on the current state and not on the sequence of events that preceded it.

In simpler terms, imagine a Markov chain as a sequence of events, where what happens next depends only on the current situation, not on how you got there.

Understanding Probabilities in a Markov Chain

In a Markov Chain, probabilities are assigned to each transition from one state to another. These probabilities are usually represented in a transition matrix if the number of states is finite. The sum of probabilities from any given state to all possible next states is always 1.

Example:
Let's consider a very simple Markov Chain with three states: A, B, and C. Suppose we have the following transitions observed:

- From A to B, 2 times.
- From A to C, 3 times.
- From B to A, 1 time.
- From B to C, 4 times.
- From C to A, 3 times.
- From C to B, 2 times.

To calculate the probabilities, we look at each state and divide the number of times each transition occurs by the total number of transitions from that state.

1. Probabilities from A:
    - To B: 2 transitions from A to B / Total transitions from A (2 to B + 3 to C) = 2 / 5 = 0.4
    - To C: 3 transitions from A to C / Total transitions from A (2 to B + 3 to C) = 3 / 5 = 0.6

2. Probabilities from B:
    - To A: 1 transition from B to A / Total transitions from B (1 to A + 4 to C) = 1 / 5 = 0.2
    - To C: 4 transitions from B to C / Total transitions from B (1 to A + 4 to C) = 4 / 5 = 0.8

3. Probabilities from C:
    - To A: 3 transitions from C to A / Total transitions from C (3 to A + 2 to B) = 3 / 5 = 0.6
    - To B: 2 transitions from C to B / Total transitions from C (3 to A + 2 to B) = 2 / 5 = 0.4

These probabilities can then be arranged in a transition matrix, which summarizes the transition probabilities from each state to every other state. For our example, the transition matrix would look like this:

	A	B	C
A	0	0.4	0.6
B	0.2	0	0.8
C	0.6	0.4	0

This matrix tells us, for example, that if we are currently in state A, there is a 40% chance that the next state will be B and a 60% chance it will be C. Similarly, the other probabilities can be interpreted from the matrix.

Markov chains are powerful tools in various fields, including statistics, economics, and computer science, due to their ability to model complex random processes in a relatively simple and understandable way.

Resources to use

- The sample code of the Markov Chain for Text Generation is given in the MarkovChain_TextGeneration_Template.py file in the drive.
- The sample training set for Text Generation of the Markov chain is given in the MarkovChain_TrainingText.txt file in the drive.

Expected Output

- A transition matrix showing the probability of transitioning from one state to another.
- A generated sequence of states starting from a randomly chosen initial state based on its frequency.

Queries

If you have any doubts/errors regarding the implementation or submission of this problem, email us at odyssey.lab@outlook.com.

Submission

Deadline: 12:00pm, 29th December, 2023
Submission Link: Google Form
You need to submit:

- The python code, e.g, MarkovChain_TextGeneration_Template.py
- A screenshot of the output of your code
- A screenshot of the transition matrix