page2image568

Reinforcement Learning

Lesson-End Project Problem Statement



**Solve a Multi-Armed Bandit Problem with the UCB Method**

**Problem Statement:**

A marketing team sends promotional messages to customers. Whenever the customers receive the promotional message, their reactions vary vastly. For every customer’s reaction, the marketing team gets a reward. You are given a set of data from which you have to find the highest rewarding message.

**Objective:** Find the message with the highest reward using UCB method.

**Dataset:**

**Link of the dataset:** <https://www.dropbox.com/s/rk376pb57fuvrmz/dataset.csv?dl=0>

**States:** sms\_1, sms\_2, sms\_3, and sms\_4 are our different states.

**Actions:** Once the customers get the messages, they can react to it in different ways, such as they can click on the link in the message, they can save the message to read it later, never read the message, they can forward the message to other people or customers, etc.

**Rewards:** For each action a reward is generated. Every number in the dataset represents a reward for taking an action to reach the successive state. For example:

|  |  |
| --- | --- |
| **Action** | **Reward** |
| Customer clicks on the link in the message | 20 |
| Customer saves the message to read it later | 10 |
| Customer never reads the message | -10 |

Higher the magnitude, higher the reward and vice versa.

**Domain:** Digital Marketing

**Prerequisites:**

* **Python**
* **NumPy**
* **Matplotlib**
* **Seaborn**

**Steps to perform:**

1. Import libraries
2. Generate the data
3. Visualize the average reward for each message
4. Initialize variables to implement UCB method
5. Implement UCB method
6. Check the reward by choosing a message after each round randomly and use UCB algorithm
7. Plot each action and the number of times each action was selected, i.e., plot messages and rewards for the reaction from the customers for the messages