Reinforcement learning for Flight Ticket Pricing (CSE 400 at 03/11/2020)

Md. Abu Shahan (170103020017 – shahanahmed668@gmail.com); Nabil Islam (170103020060 – nabilislam0047@gmail.com); (Supervisor: SC)

Objective

- We prepared a model to predict the optimal time to buy air tickets.
- At any point of time, buyers have the option to buy or wait and they are looking to purchase tickets at a cheaper price because the price of the air tickets are often fluctuating before the departure time.
- Usually, when the buyers buy air tickets, they often want to buy tickets as early as possible because they think that as sooner they buy tickets as much money they can save. But this is not reality. And for this tendency, most of the passengers buy tickets at a much higher price. Using a good model, the user can buy a ticket at a relatively low price.

Data

During this Covid-19 pandemic, the flights are limited so that we used two existing datasets to implement our model.

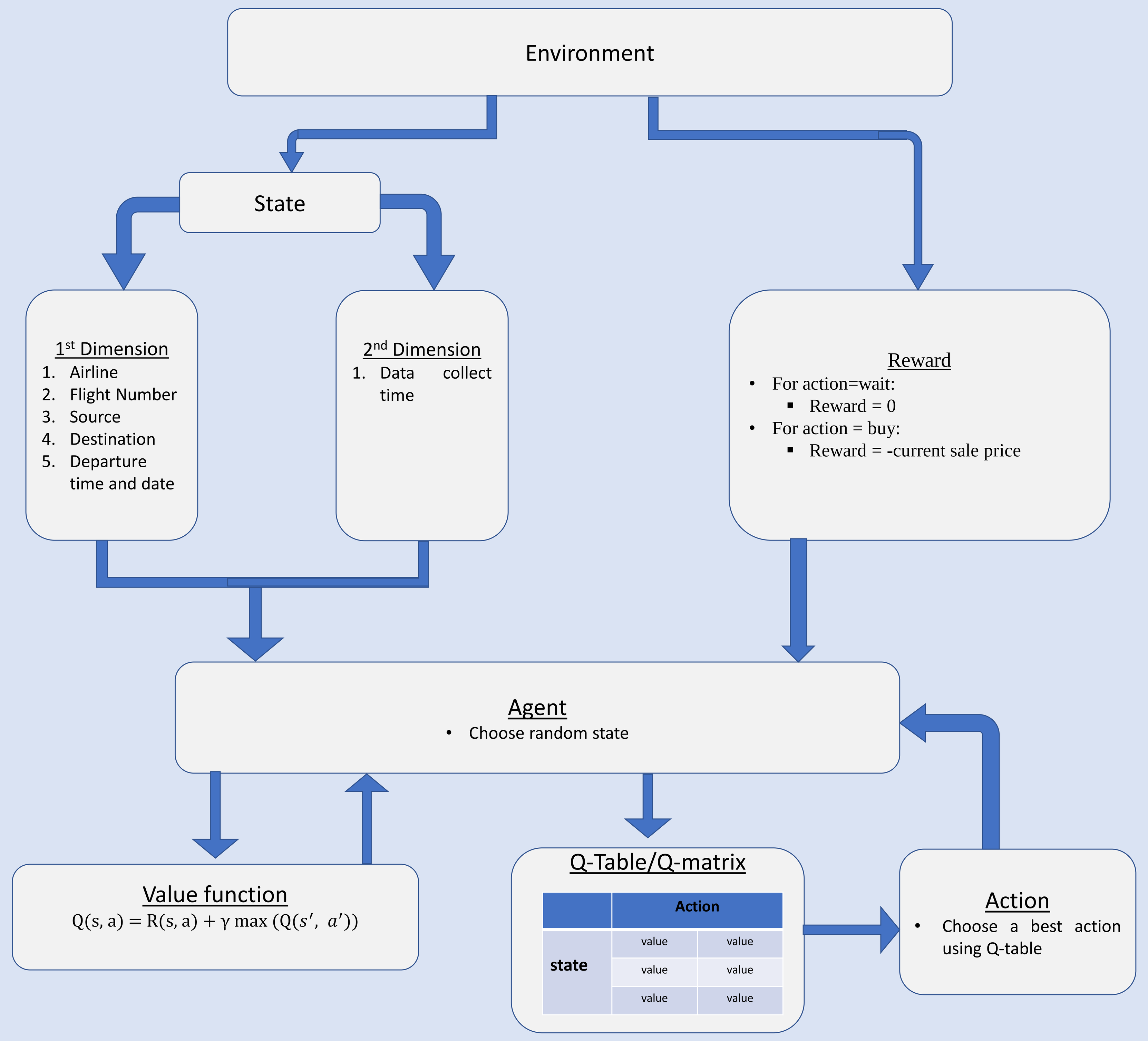
- Our first dataset contains over 12,000 fare observations for a 41 day period for six different airlines for the route Los Angeles (LAX) to Boston (BOS) and Seattle (SEA) to Washington, DC (IAD). For each departure date, the data was collected 21 days in advance at three-hour intervals.

- There are over 10,000 observations of 12 different airlines of 6 different domestic routes in India. Those routes are as follows. Bangalore - New Delhi, Bangalore – Delhi, Kolkata Our second dataset contains over– Bangalore, Delhi – Cochin, Chennai – Kolkata, Mumbai - Hyderabad

State

We divide the state into two parts. The first part has Airline, Flight Number, Source, Destination, Departure time, Departure date. And the second part has Data Collection Time.

Architecture of our model



Q-Table Creation

- Our Q-Table has three dimension.
  1. First part of the state. [Airline, Flight Number, Source, Destination, Departure Time, Departure Date]
  2. Second part of the state. [Data Collection Time]
  3. Action. [Buy, Wait]

  Example:

  Q[‘United airline, 192, LAX, BOS, 03’][‘05:00’][‘Buy’]

- First we initialized all the values of the Q-Table by zero.

Value Function

$$Q(s, a) = R(s, a) + \gamma \max (Q(s', a'))$$

- Where, s means a particular state.
- a means an action of the state s.
- Q(s, a) means a particular state of Q-Table.
- R(s, a) means the reward value of a particular state.
- γ means the discount factor.
- $\max(Q(s', a'))$ maximum value of the possible state after an action.

Our model(Q-Learning)

- First of all our agent choose a random state from the environment.
- For every action there is a reward.
- There is a discount factor γ = 1.0, this has an effect of valuing rewards received earlier higher than the rewards received later.
- Then the agent uses the value function to update the Q-Table. Q-Table values will be updated up to some specific episodes.
- The agent checks when the q-table values are maximum then show this time for buying the ticket. And when the q-table values are not maximum then show this time to wait. The decision depends on the maximum value of the q-table.

Result

Input: Airline, Flight number, Source, Destination, Departure date, Departure time, Searching date, Searching time.

Output: Buy, Wait

Example:

1. **Input:** United airline, 192, LAX, BOS, 07/11/2020, 05:00, 03/11/2020, 02:00
   Output: Wait
2. **Input:** United airline, 192, LAX, BOS, 07/11/2020, 05:00, 04/11/2020, 10:00
   Output: Buy

As you can see in the first and second example, all the inputs are similar except the Searching date and the Searching time. The agent predicts a better time when to buy.

What we have done?

- We collect two dataset to implement our model.
- We implement Q-Learning on those two dataset.
- Our model can predict the optimal time to buy the flight tickets at a relatively low price or wait for a better price in the future.

Future Work

The next time we will implement some other advanced reinforcement learning algorithms like Deep Q-Network. We think those advanced algorithms will yield better output for airline price prediction.

Conclusion

- We are able to implement Q-learning algorithm.
- This is the fact that we had more data and a larger number of features.
- The airline pricing industry has become more sophisticated, so Q-learning can no longer capture the complexity of it.

Contribution

- Most of the theoretical part(background study) done by Nabil Islam.
- Most of the implementation part done by Md. Abu Shahan.
- We collect dataset together.
- We both understand every single line of both the implementation part and the theoretical part.