I just follow the template and fill the missing segments according to the knowledge professer teach in class.

Q-learning:

First, I initialize the q table and complete the choose_action function, when this model start, the action would be taken in random, then we can choose later actions by the states we get.

Second, complete the training part, find the maximum next state action value from q-table and update current state-action pair table value by q-learning update equation

The difficulties are use execute .py, because I used to execute .ipynb on colab.
But this assign only can run on python, so I spend lots of time making this file can be execute, such as install packages.

And I also trained this model for hours, because I can't evaluate the hype parameter correctly. In the end, I searched the hint on internet for help.