Rainforcement Learning являє собою метод машинного навчання, при якому агент отримує певну винагороду на наступному тимчасовому кроці для оцінки своєї попередньої дії.

Сам метод розгортається навколо агента та середовише, в якому може і діяти агент (гра). Сам метод включає те, що поточний стан гри передається агенту, який потім, на підставі своїх значень, робить своє рішення у відповідь на цей стан гри. Потім знову агент отримує новий стан гри і вже винагороду Тому свої знання агент може оновити нагородою, поверненою самою грою. Повторюється даний алгоритм циклічно. Тобто є певні дії агента, які він може передати середовищу, поточний стан середовища і нагорода, що повертається середовищем як оцінка останнього рішення.

Хоча Q-learning алгоритм дуже потужний алгоритм, його головна слабкість – відсутність можливості знань попередніх дій. Це вказує на те, що для станів, які агент не бачив раніше, він не знає, яку дію зробити. Інакше кажучи, агент немає можливості оцінювати значення для невидимих станів. Щоб впоратися з цією проблемою, існує бібліотека DQN одна з імплементації Rainforcement Learning, де в основі лежить нейронна мережа для оцінки значень Q-функції. Навчається вона по принципу того, що на вхід мережі подаються поточні кадри ігрового поля, а виходом – відповідне значення Q для кожної можливої дії.

Також особливістю є те, що існує запис досвіду, тобто знання щодо рішень накопичуються і оновлюються. Тренування відбуваються по принципу проб і помилок. Тобто за правильність рішення, яке є кращим від попереднього надаються нагороди. Саме таким чином нейронна мережа може навчатись та робити наступні кроки на основі вже отриманого досвіду.