# Half Field Offense – Q-agent V5

In this experiment I introduced **3 major changes**. Discussed in the previous meeting First I introduced "**fixed starting points"**, which allows the agent to start playing in fixed points, in different field regions. Second change was to use **"continuous actions"**, this allows the agent to move for greater distances. Last one, I changed the agents kick speed, which **forces the agent to move near the goal**, in order to score.

This changes proved to bring a **major benefit to the player**.

## 1. Fixed Starting Points

In the beginning of each game the agent moves randomly to one of six fixed positions, which are in six different field regions ("Top Left", "Top Right", "Mid Left", "Mid Right", "Bottom Left", "Bottom Right").

This change led to a better exploration.

## 2. Continuous Actions

**Previously the agent had five actions available** ("Kick to goal", "Move/Dribble Up", "Move/Dribble Down", "Move/Dribble Left", "Move/Dribble Right"). The agent Dribbles or Moves if it has the ball or not, accordingly.

However, in order to gain speed and moves for a great distance, the agent needs to do the same action for a long time, which is very hard to do during exploration. The agent would end up alternating between different actions, and, as result of it, the agent would almost stay in the same place.

So, in order to go throw this problem I **introduced long and short movements**, the long movements makes the agent to select the same action 20 times. The short movement leds the agent to select the same action 10 times. However, while learning, for the agent it will only count as 1 action.

To sum up, the agent **now can select nine different actions** ("Kick to goal", "Long Move/Dribble Up", "Long Move/Dribble Down", "Long Move/Dribble Left", "Long Move/Dribble Right", "Short Move/Dribble Up", "Short Move/Dribble Down", "Short Move/Dribble Left", "Short Move/Dribble Right").

## 3. Weaker Kick

Previously, the agent was able to kick the ball with the maximum speed possible. Which led to scoring far way for the goal. So I reduced the speed, which forced the agent to learn to move near the goal, in order to score.

# 1. Game parameters:

- Number of **teammates: 0;**

- Number of **opponents: 1;** (Dumb Goalie)
- Number of **episodes: 10k train;**

# 2. Q-Learning Agent

## 2.1. Q learning parameters:

- Learning rate: 0.10;
- Epsilon values: [0.8, 0.7, 0.5, 0.4, 0.3];
- Discount factor: 0.99;
- **Q learning table dim:** 120 environment states * **9** number of actions;

## 2.2. State Features

1. **Position** – int [0, 6], subdivided the field in 6 regions, each integer signalizes one of these regions;
2. **Should Shoot** – int (0,1). Agent's goal opening angle > 20%;
3. **Opponent is close** – int (0,1). Opponent is close;
4. **Ball Position –** int [0, 4]. Five states {0: "Player Has Ball", 1: "Up", 2: "Right", 3: "Down", 4: "Left"}

## 2.3. Actions

1. **Kick to Goal –** Shoot to the goal. This command only works when the agent has the ball.
2. **Short Dribble/Move Up –** The agents goes up 10x. If it has the ball, it dribbles the ball up, otherwise just moves up.
3. **Short Dribble/Move Down** – The agents goes up 10x. If it has the ball, it dribbles the ball up, otherwise just moves up.
4. **Short Dribble/Move Right** – The agents goes up 10x. If it has the ball, it dribbles the ball up, otherwise just moves up.
5. **Short Dribble/Move Left** – The agents goes up 10x. If it has the ball, it dribbles the ball up, otherwise just moves up.
6. **Long Dribble/Move Up –** The agents goes up 20x. If it has the ball, it dribbles the ball up, otherwise just moves up.

7. **Long Dribble/Move Down** – The agents goes up 20x. If it has the ball, it dribbles the ball up, otherwise just moves up.
8. **Long Dribble/Move Right** – The agents goes up 20x. If it has the ball, it dribbles the ball up, otherwise just moves up.
9. **Long Dribble/Move Left** – The agents goes up 20x. If it has the ball, it dribbles the ball up, otherwise just moves up.

### 2.4. Rewards

- - 1 – the game ends without scoring goal;
- + 1 – score goal;

# 3. Goalkeeper Agent

The agent presents 2 simple behavior:

**If** ball is on the upper side of the field:

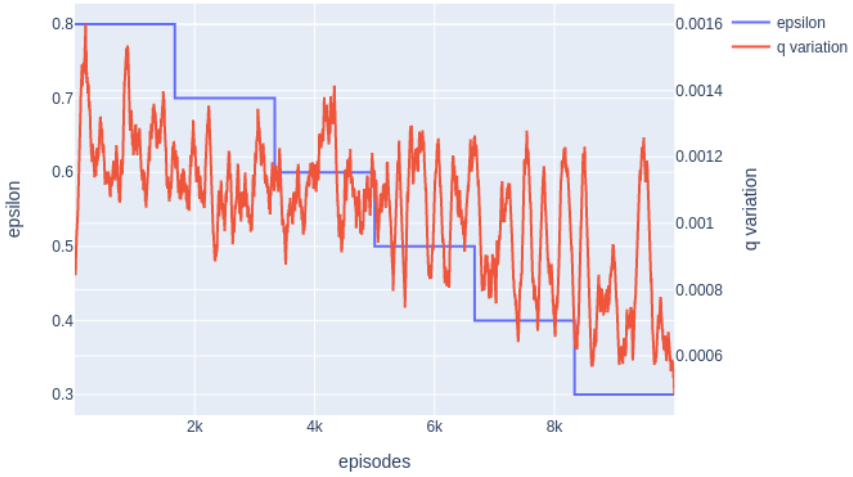
the goalkeeper moves to the goal top corner;

**Else**:

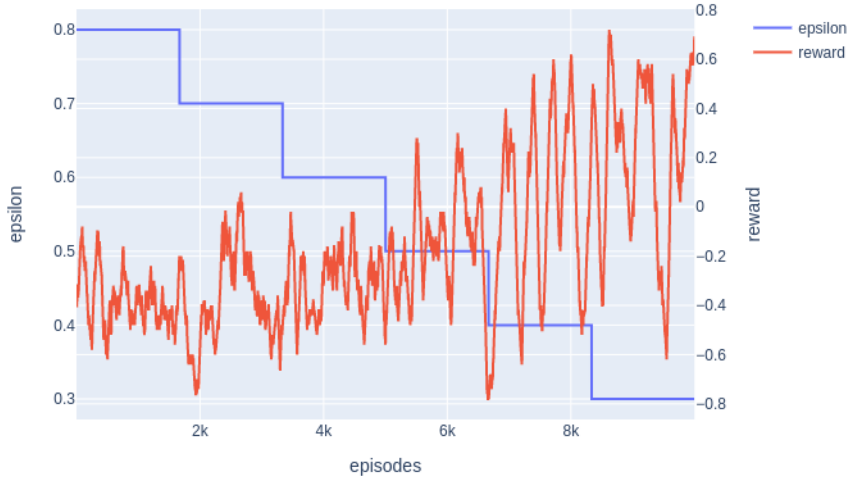
the goalkeeper moves to the goal bottom corner;

# 4. Results:

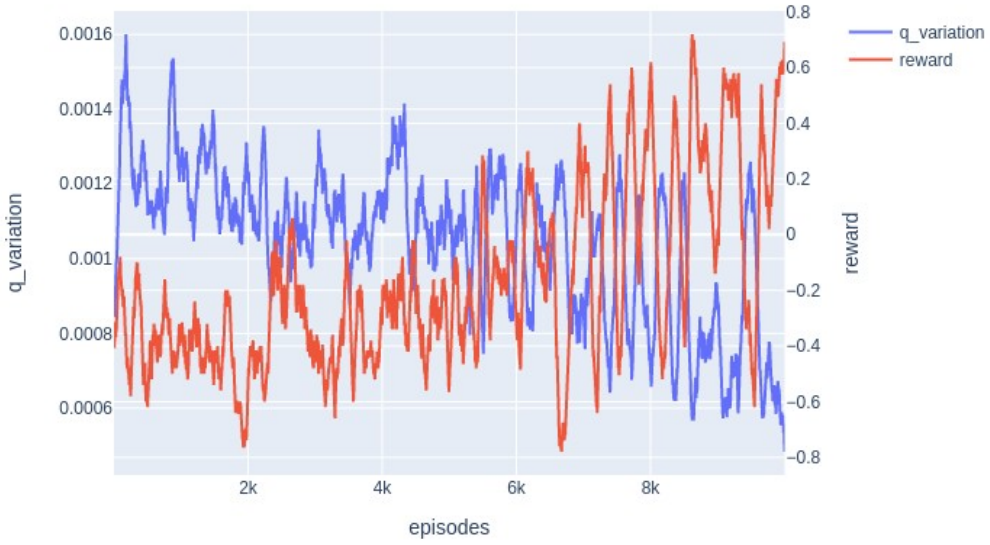
### 4.1. Relation q-learning variation – epsilon variation:



## 4.2. Relation epsilon variation – reward variation:




## 4.3. Relation q-table variation – reward variation:



## 4.4 Exploration Heat-map:

On the y-axis is the 6 different regions, on the x-axis bar are the 9different actions. This sums the total actions selected at each state, during all the training.