Bias in environments in which React

(1) Bias in Environments in which React

(2) Embodiment Problem

(3) Background Changes (Preferences, Weather)

\[ P(x, y) = P(x) \cdot P(y | x) \]

\[ \text{input} \quad \text{output} \]

\[ \text{Concept Shift} \quad \text{Label Shift} \]

Algorithm

(1) Collect Data from Human Expert

\[ D = \{(x_1, y_1^*), (x_2, y_2^*), \ldots\} \]

(2) Train \( \pi_0 \) on \( D \).

(3) Execute \( \pi_0 \)

(4) Call Interactive Expert on All States \( \pi_0 \)

\[ D = \{(x_1, y_1), (x_2, y_2), (x_3, y_3)\} \]

\[ \pi: X \rightarrow Y \]
$$\text{Avg Regret} = \frac{1}{T} \left( \sum_{t=1}^{T} l_t(\pi_t) - \min_{\pi^* \in \Pi} \sum_{t=1}^{T} l_t(\pi^*) \right)$$

For $T \to \infty$, any regret $\to 0$

No Regret

Policy in hindsight