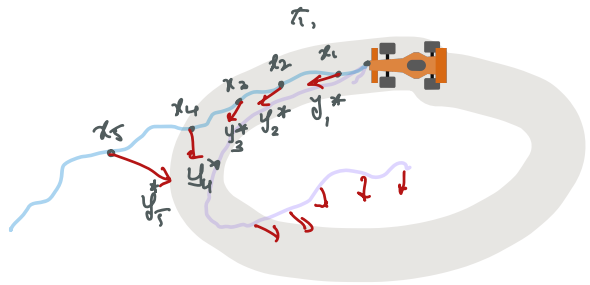


IMITATION LEARNING → INTERACTIVE LEARNING

DATASET 1 INDUCED BY  $\pi_1$

$$\{ (x_1, y_1^*), (x_2, y_2^*), \dots \}$$

↓  
LOSS FUNCTION  $l_1(\pi)$



$$l_1(\pi) = \mathbb{E}_{x \sim P(\cdot | \pi_1)} \mathbb{1}(\pi(x) \neq y^*)$$

$y^* \sim$  HUMAN DEMONSTRATION

DATASET 2

$$\{ (x_1, y_1^*), \dots \}$$

⇒  $l_2(\pi)$

\* EVERY ROUND OF ITERATION CREATES A LOSS FUNCTION

\* LOSS FUNCTIONS VARY IN COMPLEX WAYS

LOSS FUNCTIONS OVER ROUNDS

	0.2	0.7	
	0.1	$\pi_2^{1.0}$	
	$\frac{1.0}{\pi_1}$	0.3	
	0.3	0.1	
	:	:	
	.	:	

GREEDY

$$\pi_i = \operatorname{argmin}_{\pi \in \Pi} l_{i-1}(\pi)$$

Avg

$$\text{REGRET} = \frac{1}{N} \left( \sum_{i=1}^N l_i(\pi_i) - \min_{\pi \in \Pi} \sum_{i=1}^N l_i(\pi) \right)$$

(LOSS OF LEARNER) (LOSS OF BEST POLICY IN HINDSIGHT)

$\pi_1$	0	1	0	1	-	-
$\pi_2$	1	0	1	0	-	-

$$1 + 1 + 1 + 1 + \dots = 2$$

$$0 + 1 + 0 + 1 + 0 + 1 + \dots = 2$$