

$$\hat{E}[Y | S]$$

**User clicks on item**

**Features about the  
user and the item**

# Overlap in $S$ space

- If you collected and trained your model on Sample 1, you'd be a little worried about predicting on Sample 2.
- This is a common source of ML problems that we're trained to check for:
  - data drift, concept drift, logging errors, bugs, model staleness

