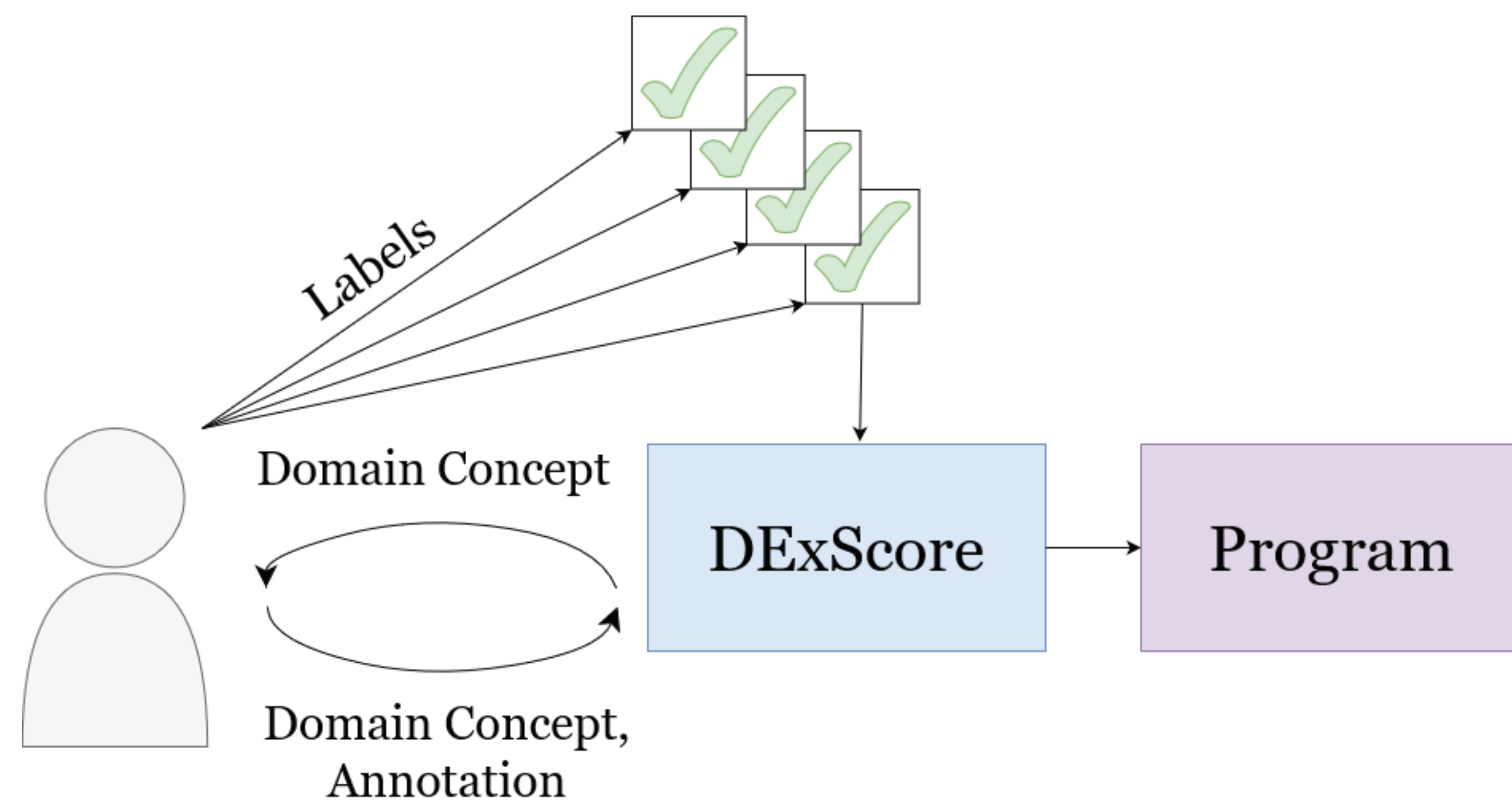


Eliciting Domain Expertise Reduces Examples Needed for Program Synthesis

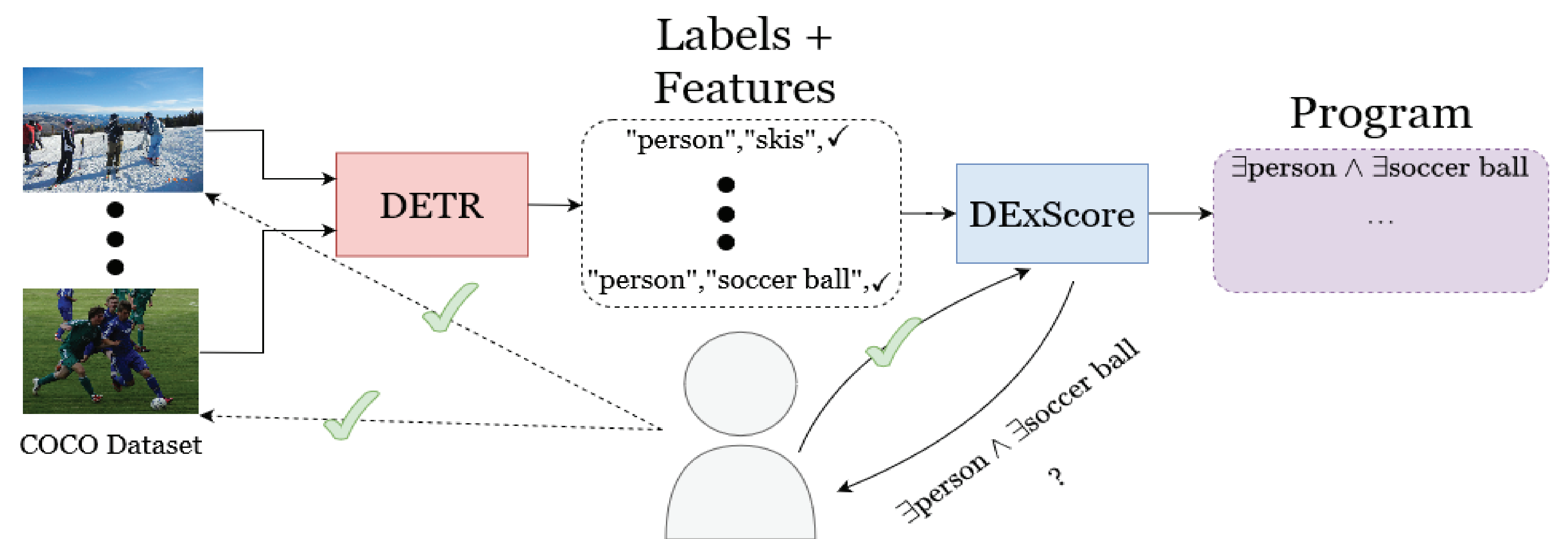
Jeremy Ferguson, Justin Lubin, Sarah E. Chasins

Framework



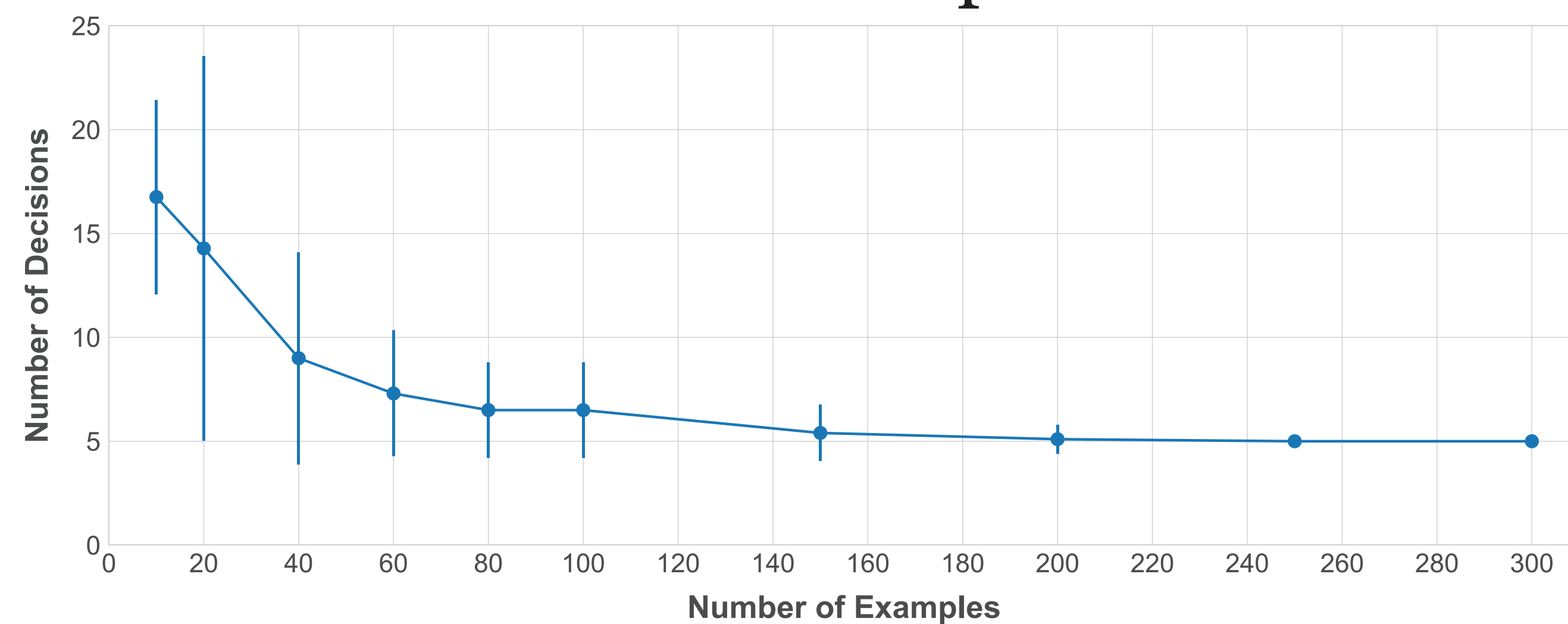
DExScore Instantiation

Sample Task: "Is this image one of people playing sports?"

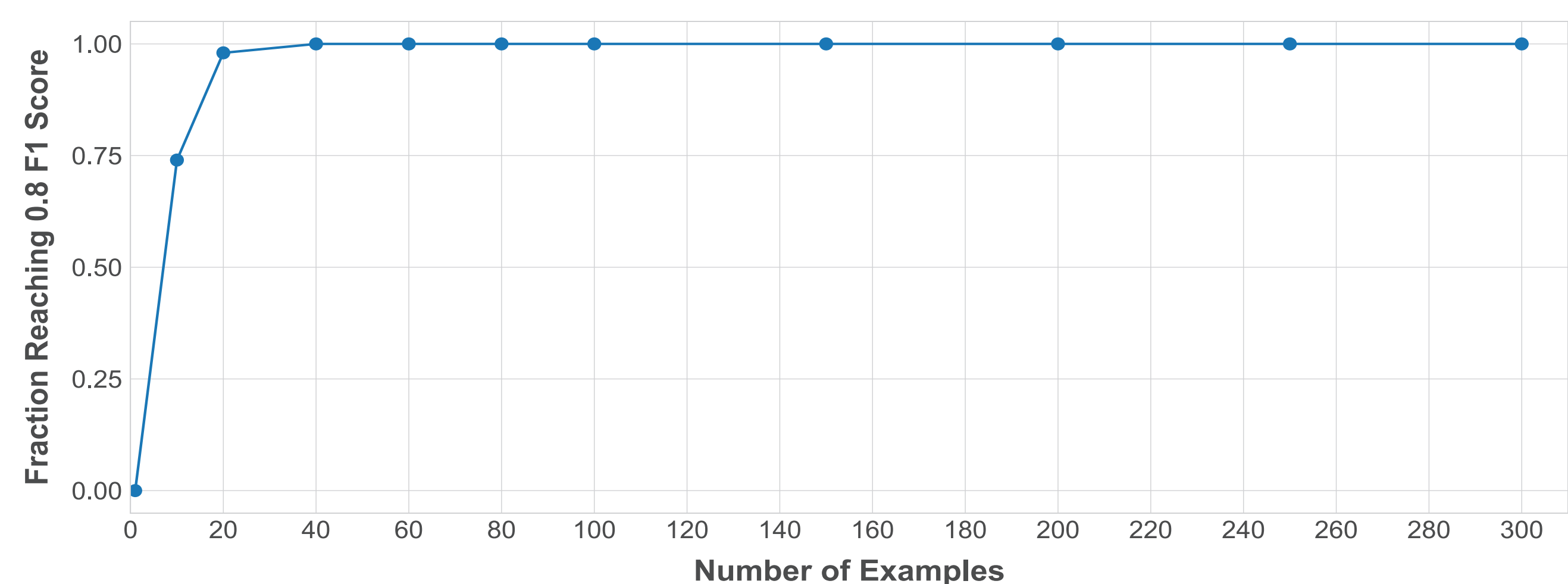


Results

Decisions-Examples Tradeoff



Fraction of Runs that Reach 0.8 F1 Score



Technique

- Use DSL of predicate functions to score the document
- How to suggest new predicates?

- Use log odds ratio to rank them:

$$\text{RATIO}(p, X^+, X^-) = \log_2 \frac{\mathbb{P}(p | X^+)}{\mathbb{P}(p | X^-)} = \log_2 \sum_{x \in X^+} \frac{|p|(x)}{|X^+|} - \log_2 \sum_{x \in X^-} \frac{|p|(x)}{|X^-|}$$

- Extractor functions to compute X^+ and X^- :

$$\text{EXTRACT}(P, X) = (X \cup \{x \in X \mid \llbracket P \rrbracket(x) > t^+\}, \{x \in X \mid \llbracket P \rrbracket(x) \leq t^-\})$$

- Mutual information filtering:

$$\text{MI}(X, \mathcal{Y}) = \sum_{x \in X} \sum_{y \in \mathcal{Y}} \mathbb{P}(x, y) \cdot \log_2 \frac{\mathbb{P}(x, y)}{\mathbb{P}(x) \cdot \mathbb{P}(y)}$$

LLM Baseline Comparison

