UNIVERSITY
OF TRENTO - Italy

# Interactive Machine Learning

**Andrea Passerini**
credits at the end..

**Spring Workshop on Mining and Learning - 2020**

# Is AI a threat to humankind?



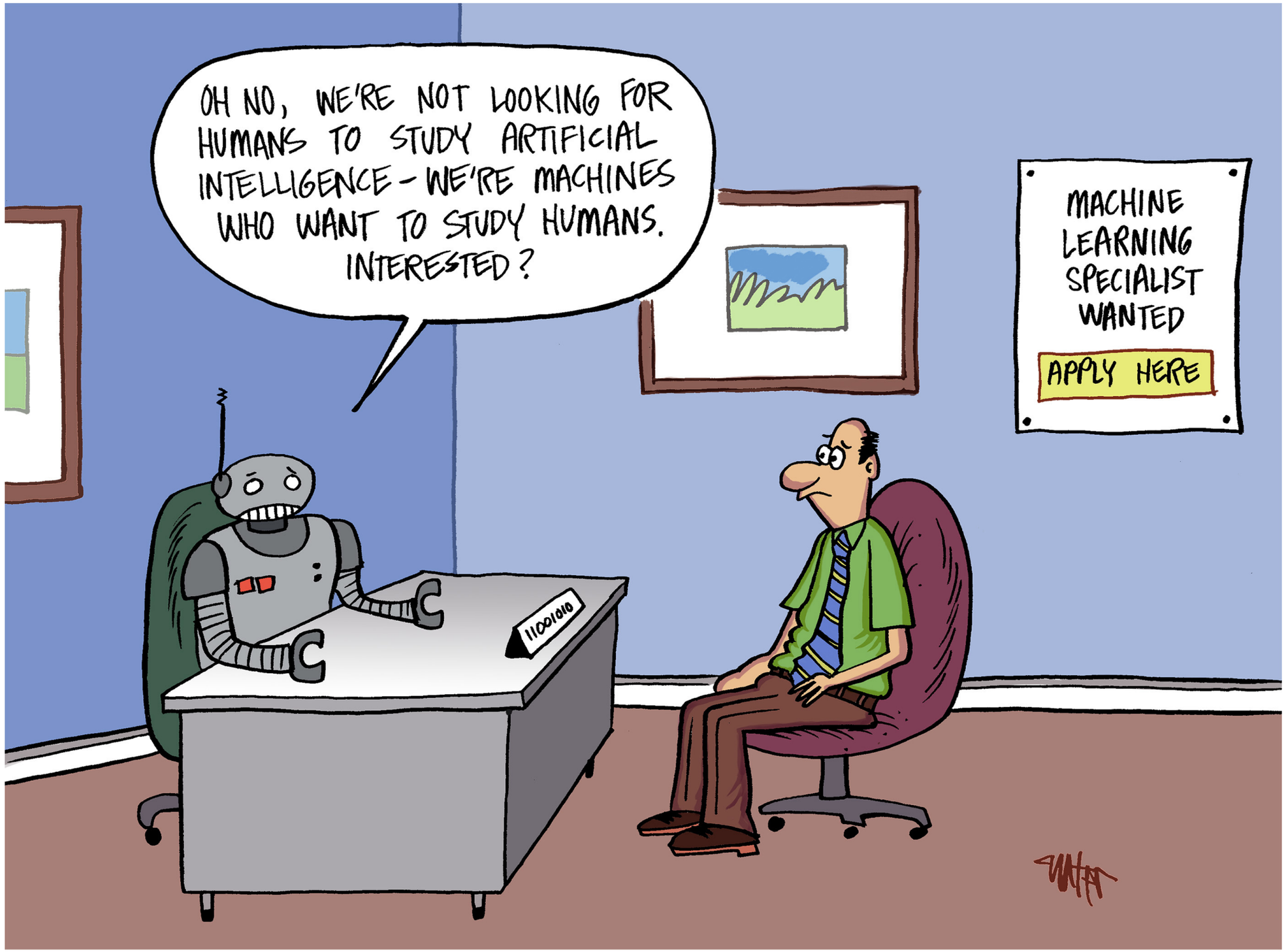
"We're looking for someone with your exact qualifications, but a mechanical version."

CartoonStock.com

# or an opportunity to augment humans?



# maybe AI needs to first understand humans!



# **Questioning** user feedback

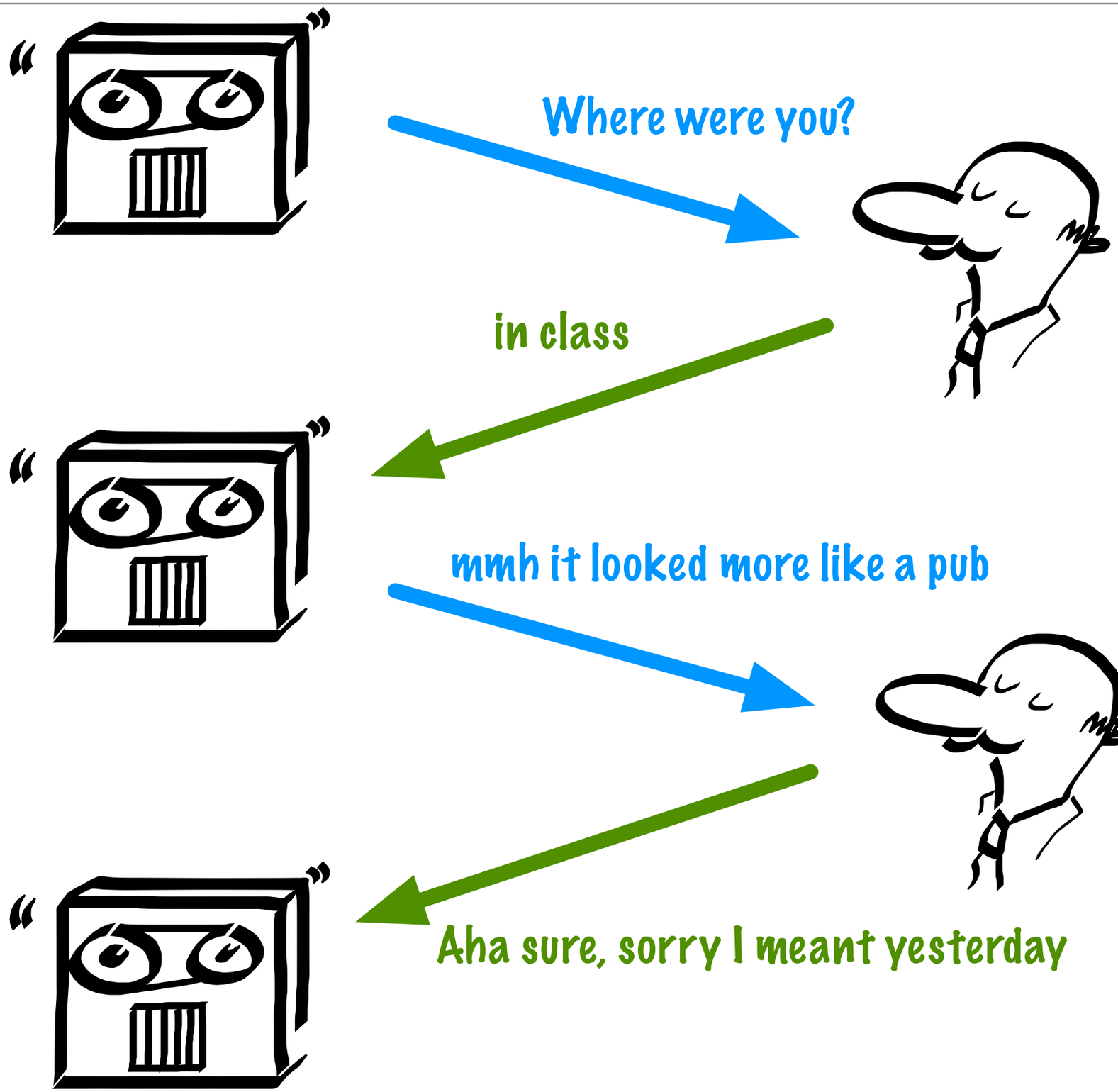
PROBLEM

- User **feedback** can be **inconsistent**
- Machine learning techniques are **only robust to moderate** amount of **noise**

SOLUTION

- Identify when user **feedback is unconvincing**
- **Discuss** with user to **solve conflict**

# Skeptical learning: interaction



# Skeptical learning: features

- **Online active learning** to decide when to query user
- **Confidence ratio** (btw machine and user label) to decide when to contradict
- **Conflict resolution** strategy (contradict user, accept second answer)

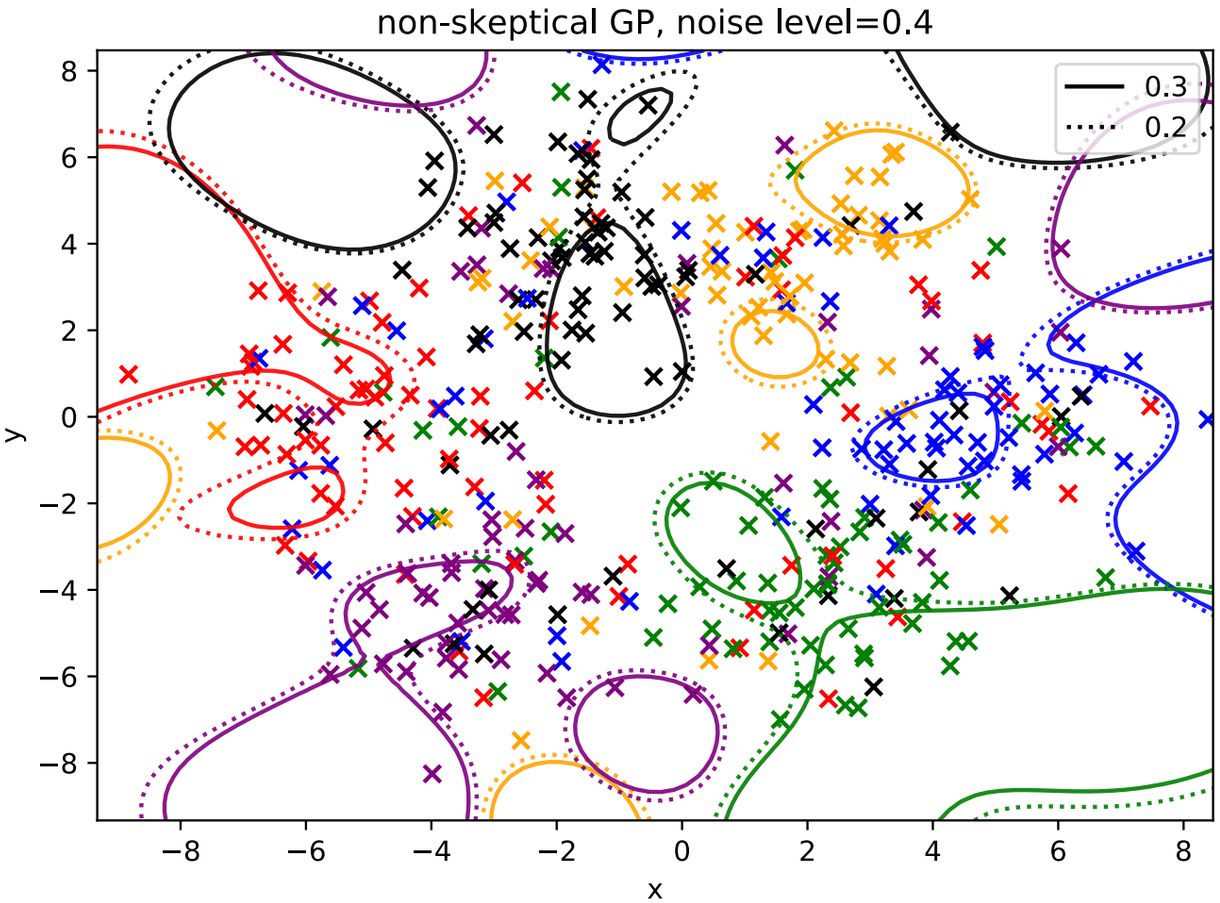
# Incremental Gaussian Skeptical Learning (ISGP)

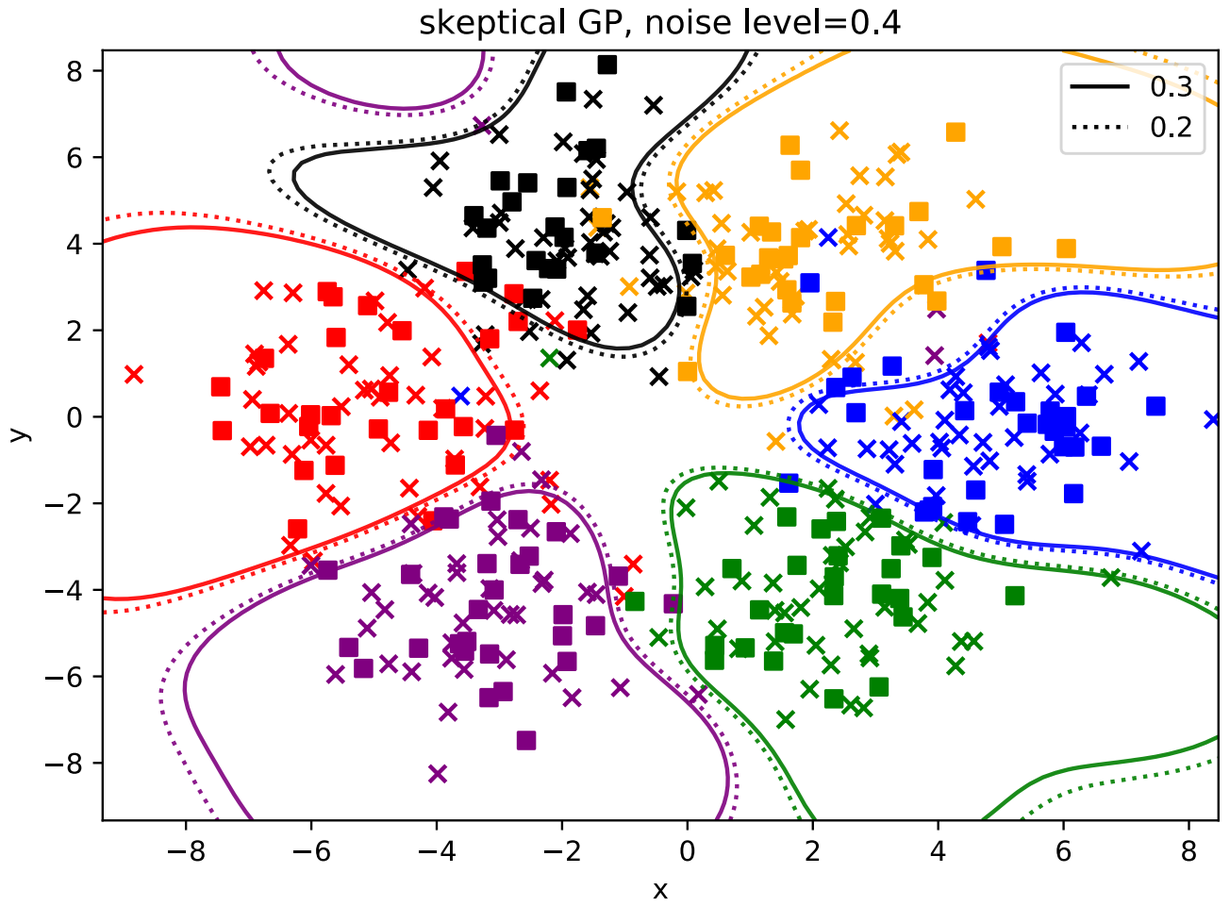
- **Gaussian process** to manage uncertainty
- **Incremental example updates** to scale up
- **Incremental label update** to deal with novel classes

# Skeptical learning: learning algorithm

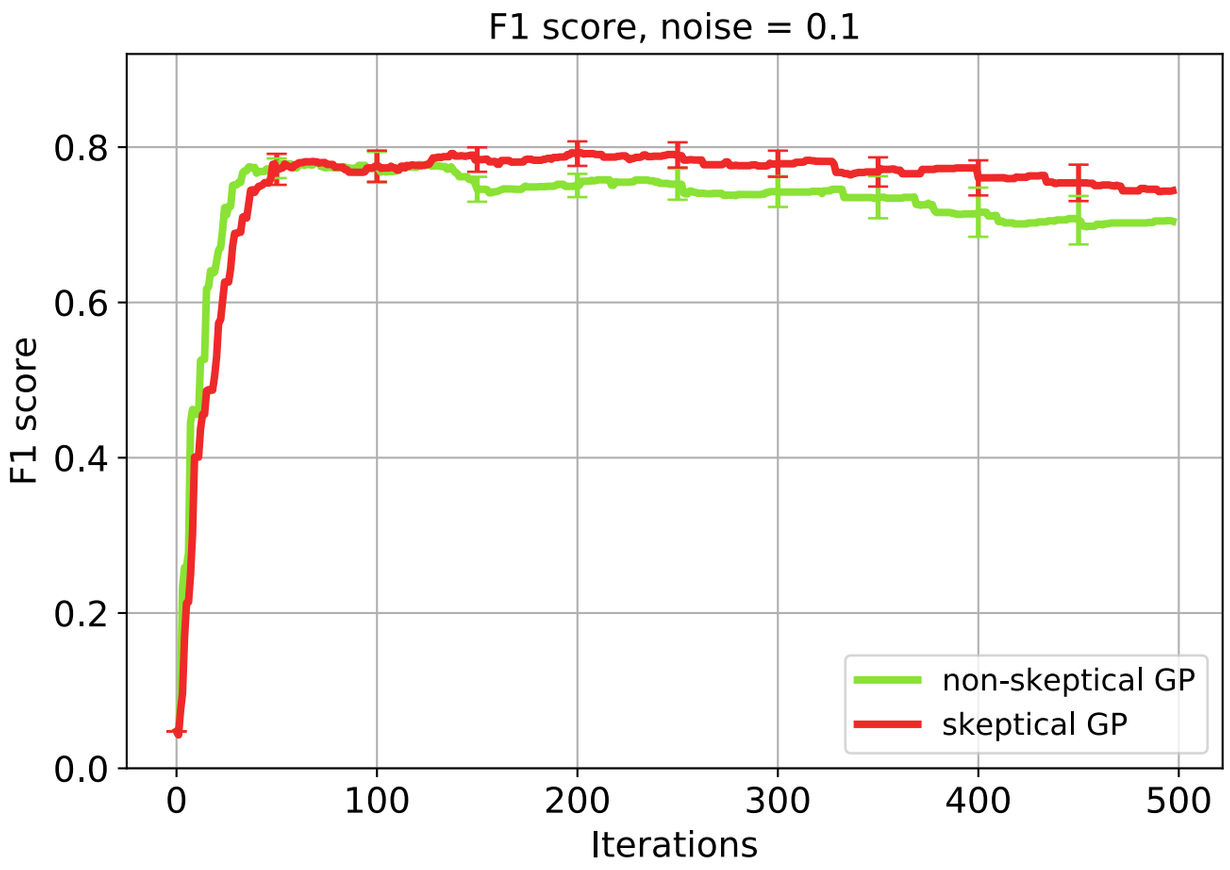
**for** $t = 1, 2, \ldots$ **do**
  receive $x_t$
  $\hat{y}_t \leftarrow \operatorname{argmax}_{y \in \mathcal{Y}_{t-1}} \ \mu_y(x_t)$
  **if** uncertain about $\hat{y}_t$ **then**
    request label, receive $\tilde{y}_t$
    **if** skeptical about $\tilde{y}_t$ **then**
      challenge user with $\hat{y}_t$, receive $y'_t$
    **else**
      $y'_t \leftarrow \tilde{y}_t$
    update the GP with $(x_t, y'_t)$
    $\mathcal{Y}_t \leftarrow \mathcal{Y}_{t-1} \cup \{y'_t\}$

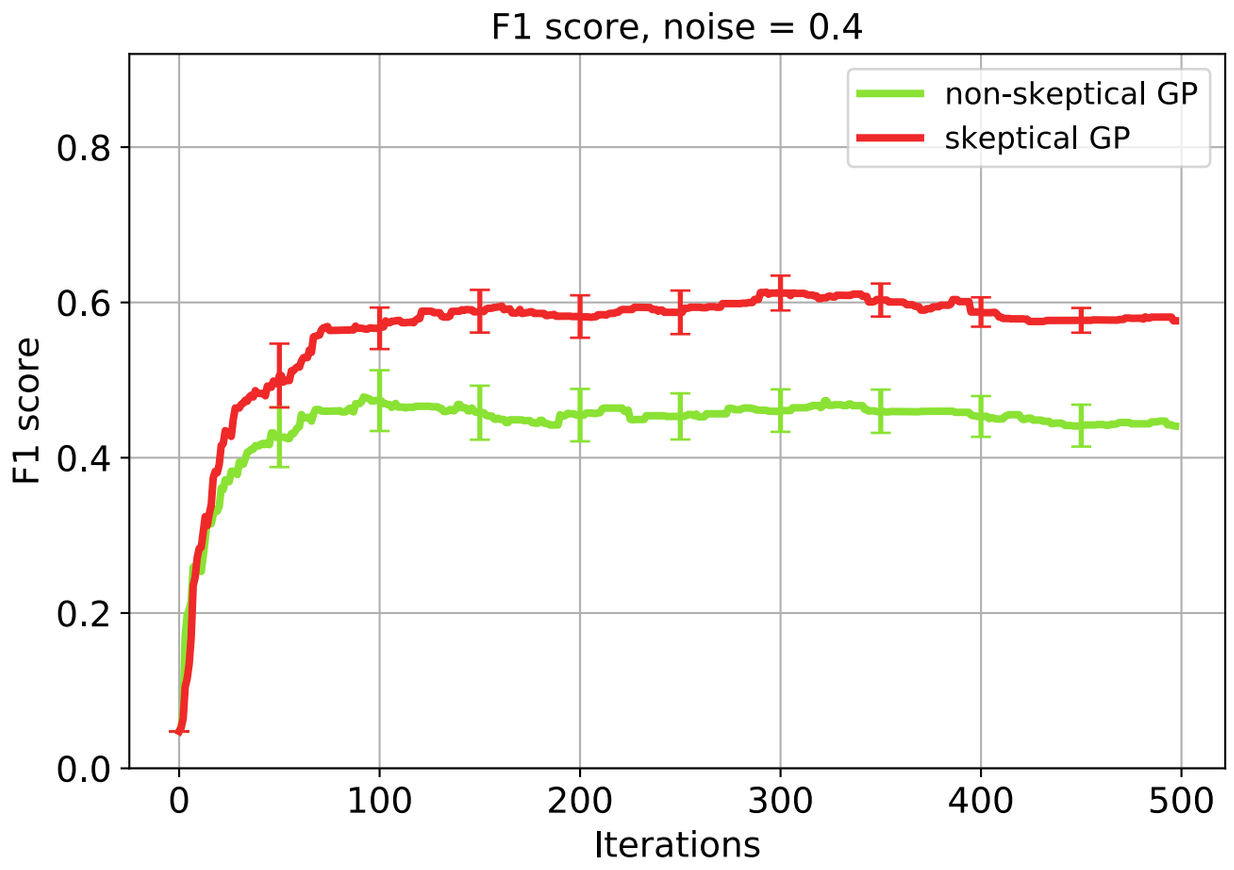
# Skeptical learning: robustness



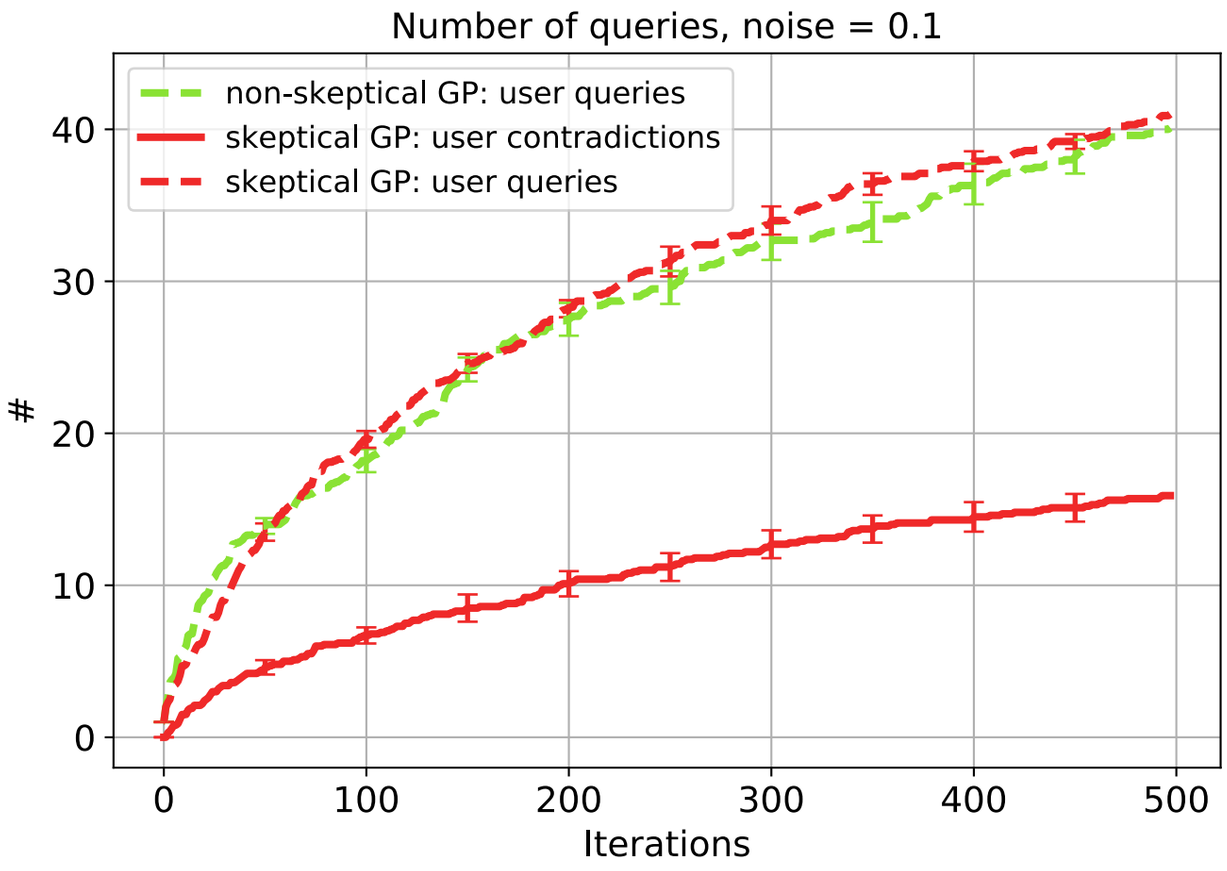



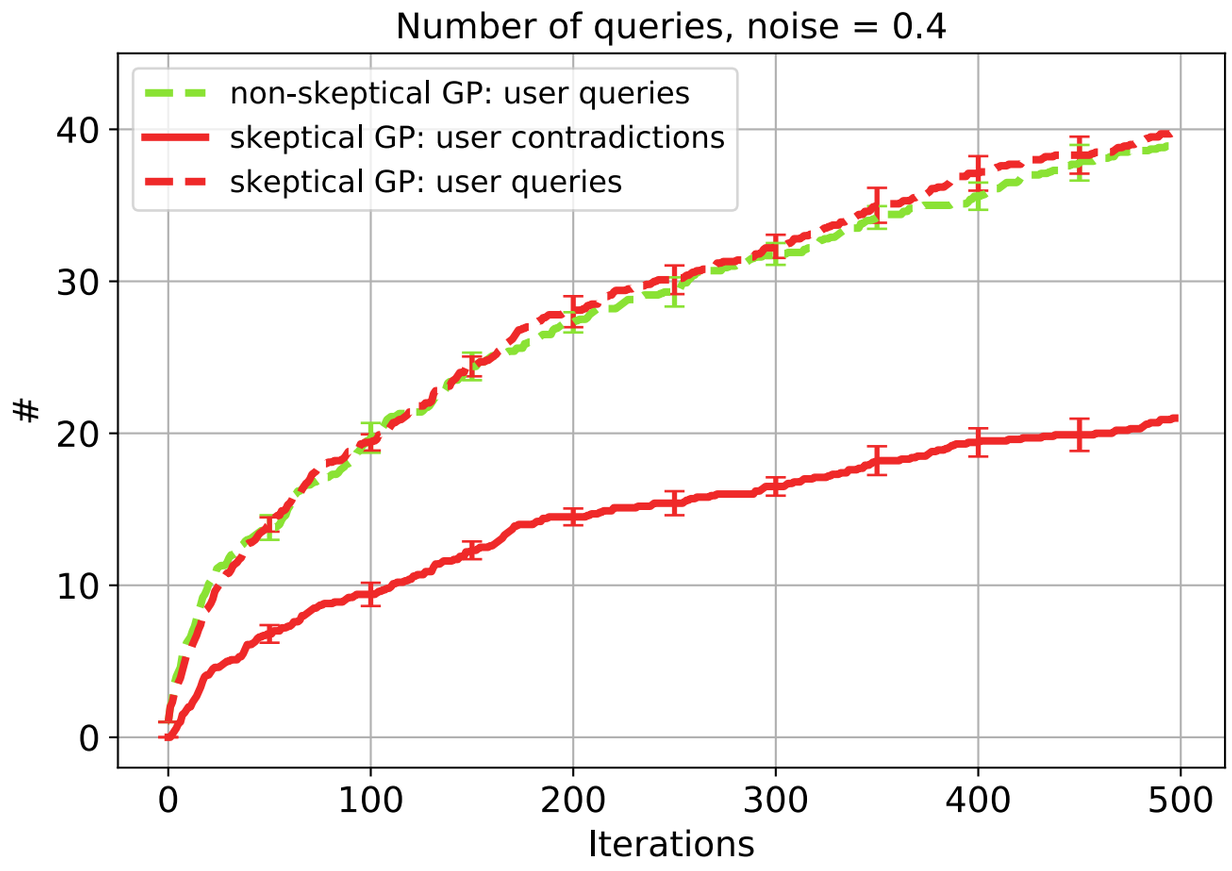
# Skeptical learning: predictive performance



# Skeptical learning: # of queries



# Adapting to **user perspective**

PROBLEM

- Global learning of concepts has problems in progressively **incorporating new concepts**
- Global learning of concepts **ignores subjectivity** in concept definition (e.g. my cat vs a siamese cat)

# Adapting to **user perspective**

SOLUTION

- Focus on **egocentric viewpoint** on concepts
- Use **similarity** as driving principle for instance-level recognition
- Use **persistency** for object identification from “few samples”
- Perform **developmental learning** by progressively introducing new objects

# Egocentric viewpoint

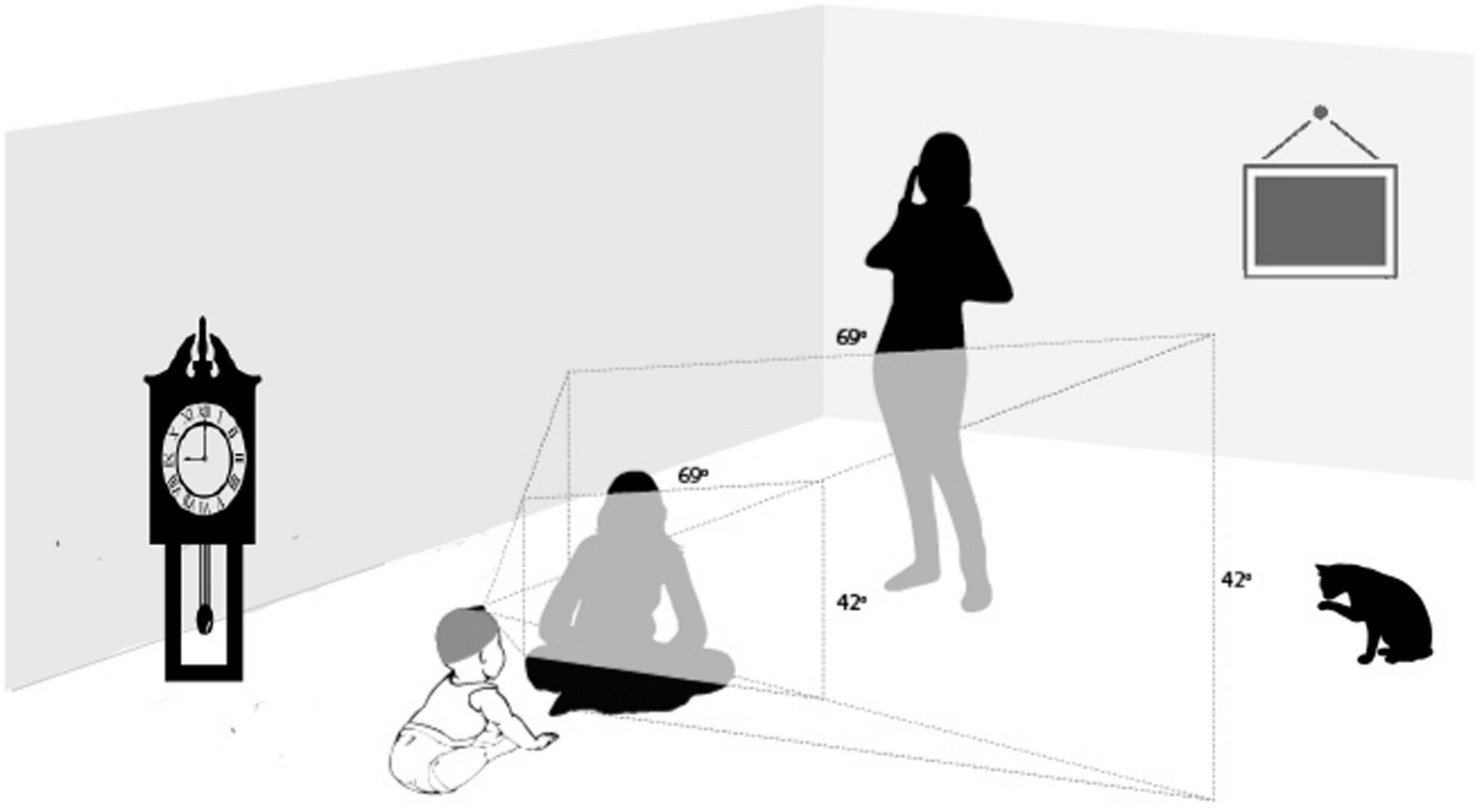



image taken from Smith & Slone, *A Developmental Approach to Machine Learning?* Frontiers in Psychology, 2017

# Similarity-based learning





- Focus on **instance-level** recognition (highest granularity)
- Use **perceptual similarity** as learning primitive
- Model higher-level **concepts by aggregation**

# Persistency-based learning

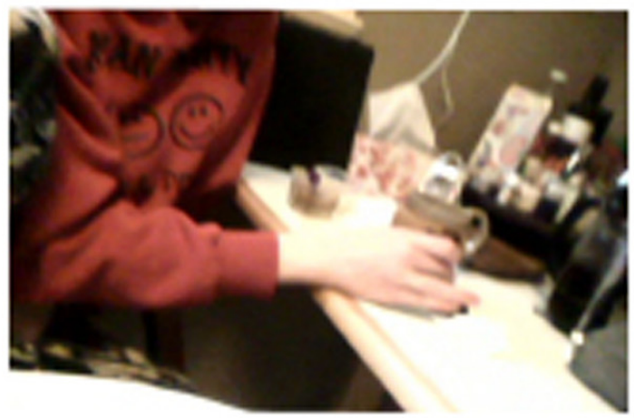




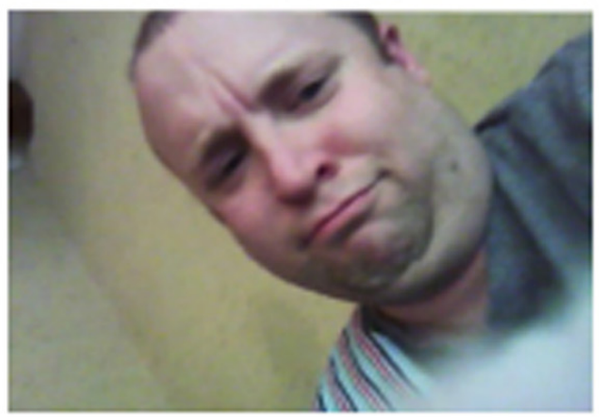
- Acquire instances using videos
- Focus video on instance, show multiple viewpoints
- Model visual invariances

# Developmental learning













1 to 3 months

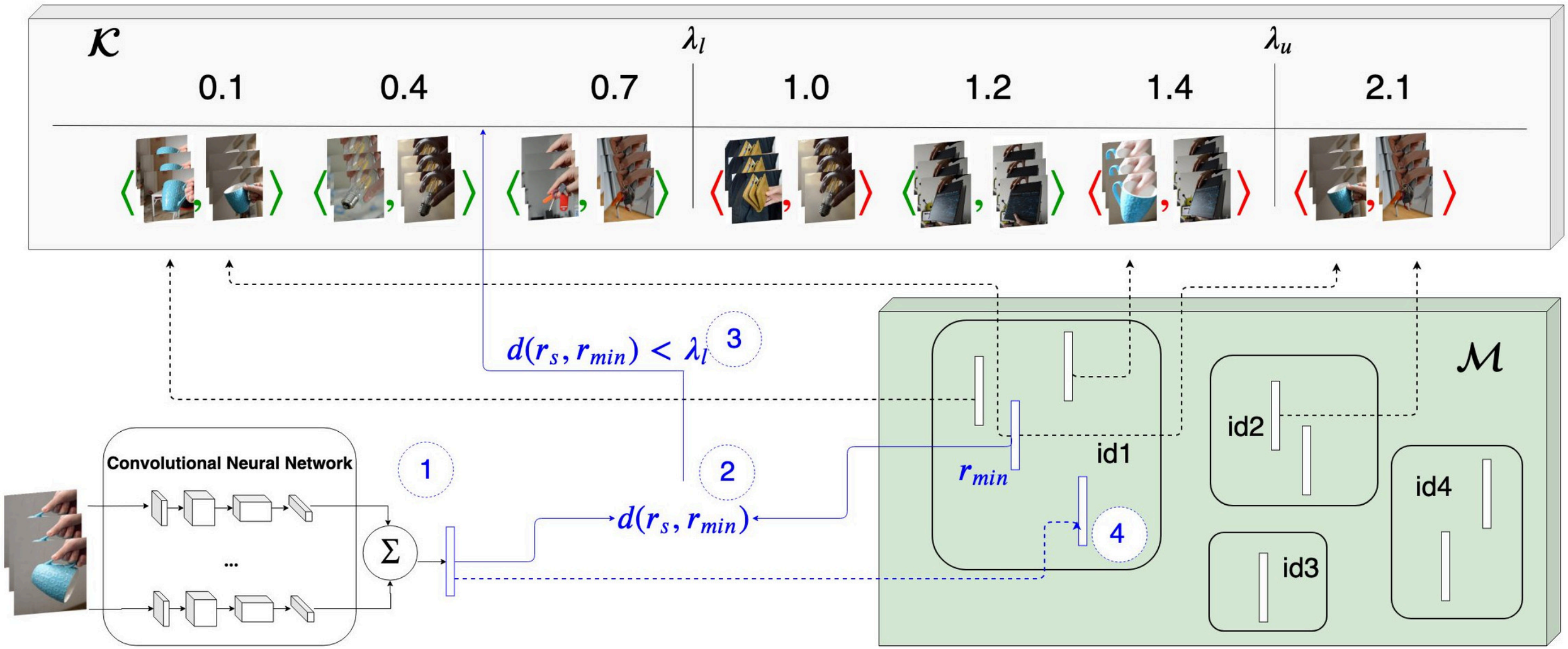
8 to 10 months

12 to 24 months

- Incrementally introduce new objects
- Progressively reduce supervision

image taken from: ibid

# Continual egocentric object recognition




**FOLLOWER**

# Follower: selection of thresholds

entropy

$$\underset{\lambda_l,\lambda_u}{\text{argmax}} \quad H(\mathcal{Y}_{\lambda_l}^{\lambda_u}) - H(\mathcal{Y}^{\lambda_l}) - H(\mathcal{Y}_{\lambda_u})$$

pairwise label

$$\text{subject to:} \quad \mathcal{Y}_{\lambda_l}^{\lambda_u} = \{y | \langle y, \delta \rangle \in \mathcal{K} \wedge \lambda_l \leq \delta \leq \lambda_u\}$$

$$\mathcal{Y}^{\lambda_l} = \{y | \langle y, \delta \rangle \in \mathcal{K} \wedge \delta \leq \lambda_l\}$$

$$\mathcal{Y}_{\lambda_u} = \{y | \langle y, \delta \rangle \in \mathcal{K} \wedge \lambda_u \leq \delta\}$$

feedback memory

$$\sum_{i=1}^{|\mathcal{K}|} \mathbb{1}(\lambda_l \leq \delta_i \leq \lambda_u) = \lceil \alpha |\mathcal{K}| \rceil$$

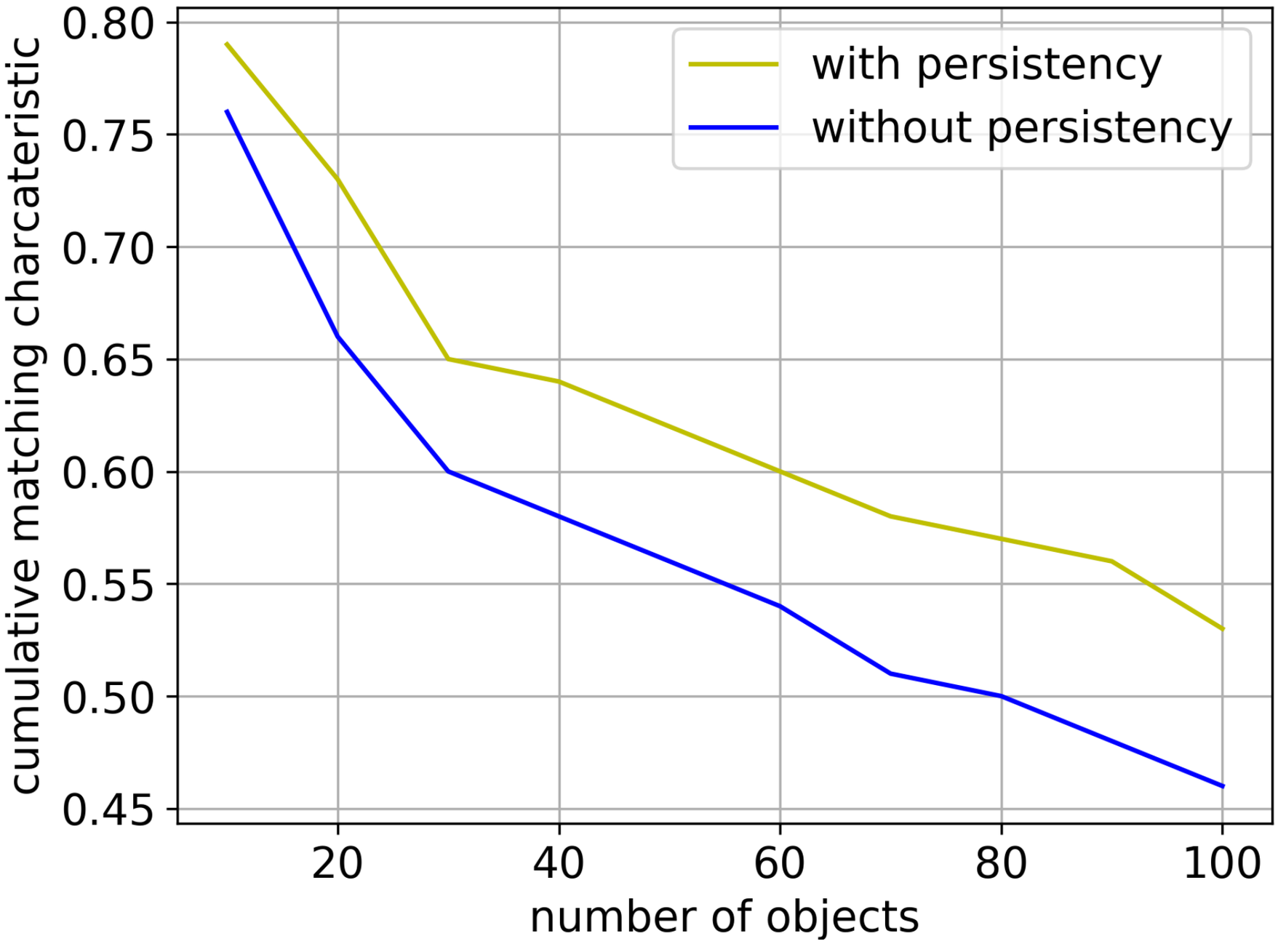
**is same** threshold

pairwise distance

**is different** threshold

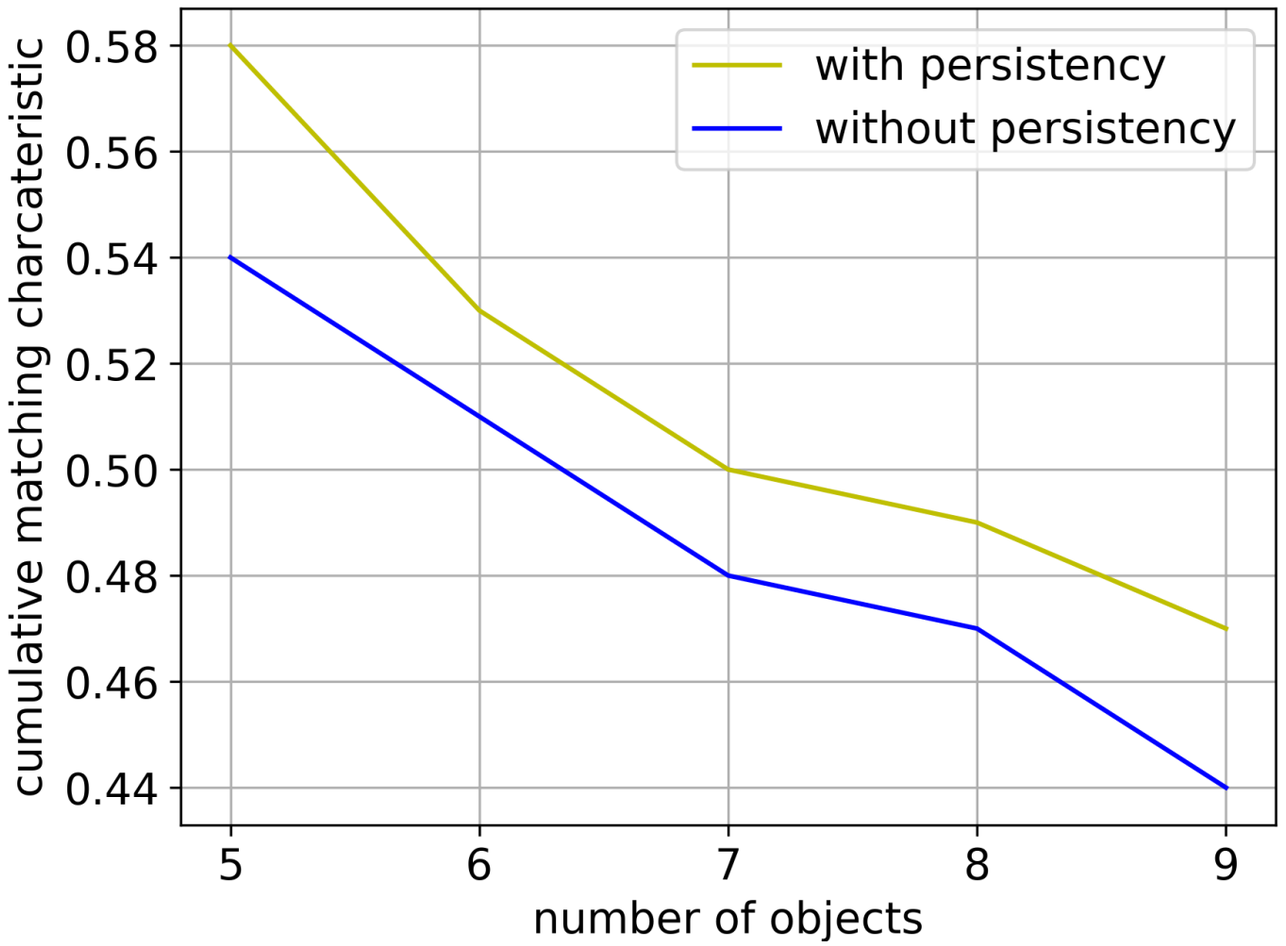
expected % of queries

# Experiments: role of persistency



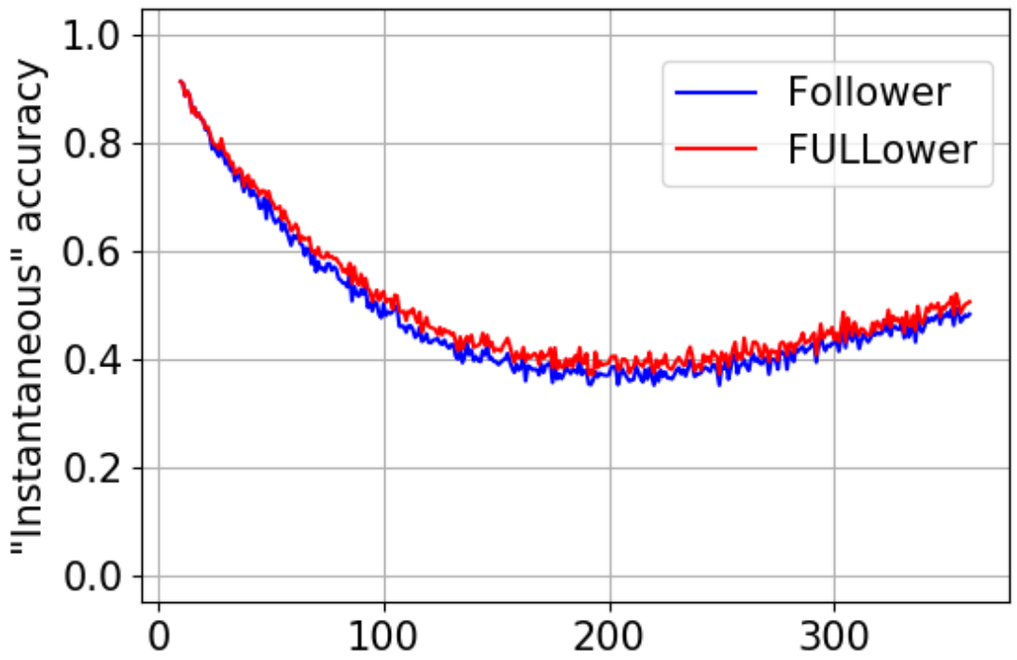
# Experiments: instance-level recognition

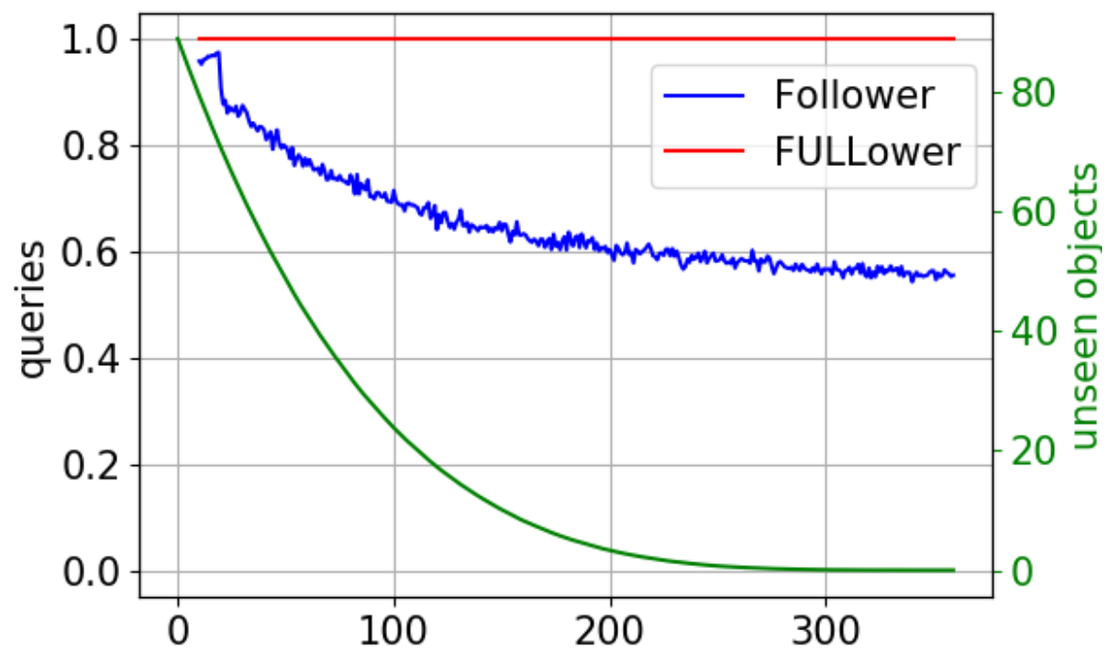




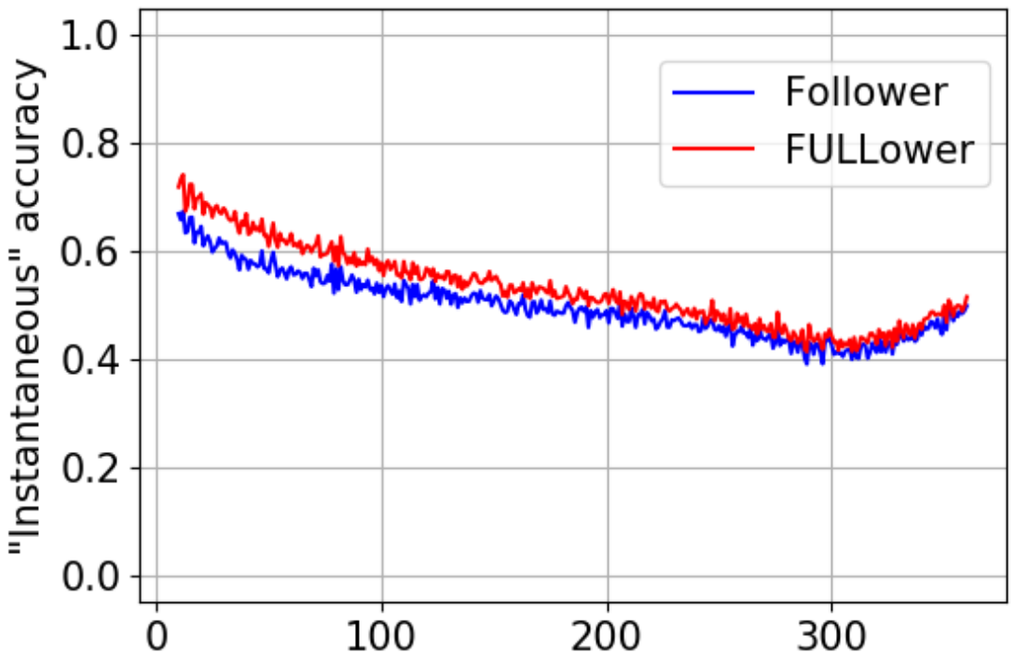
# Experiments: role of developmental learning

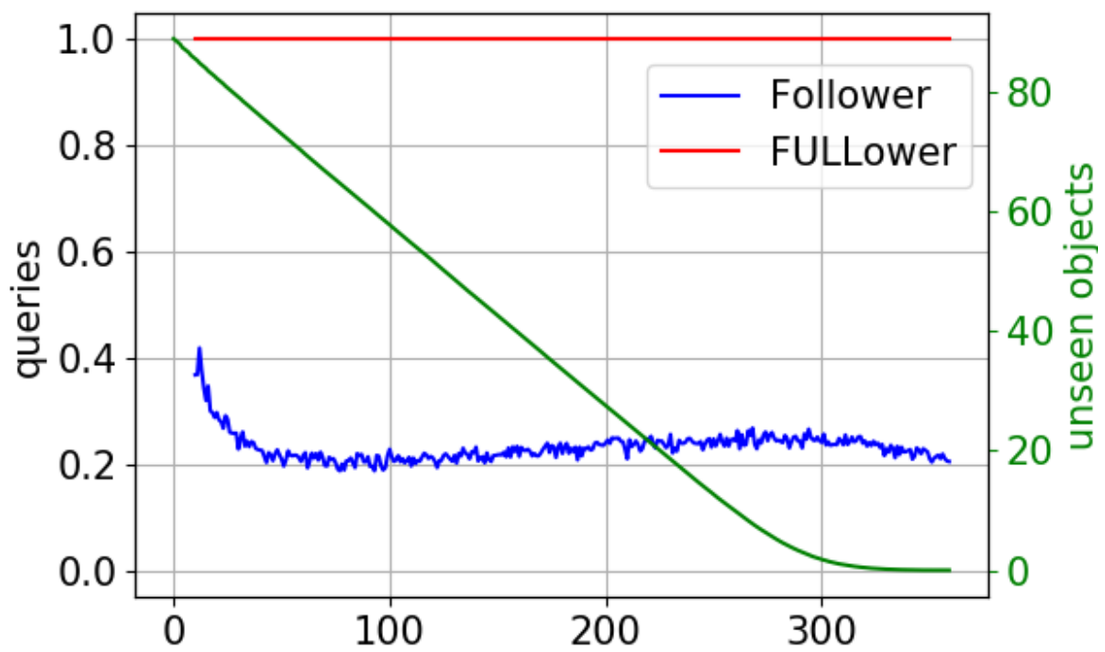
## Random



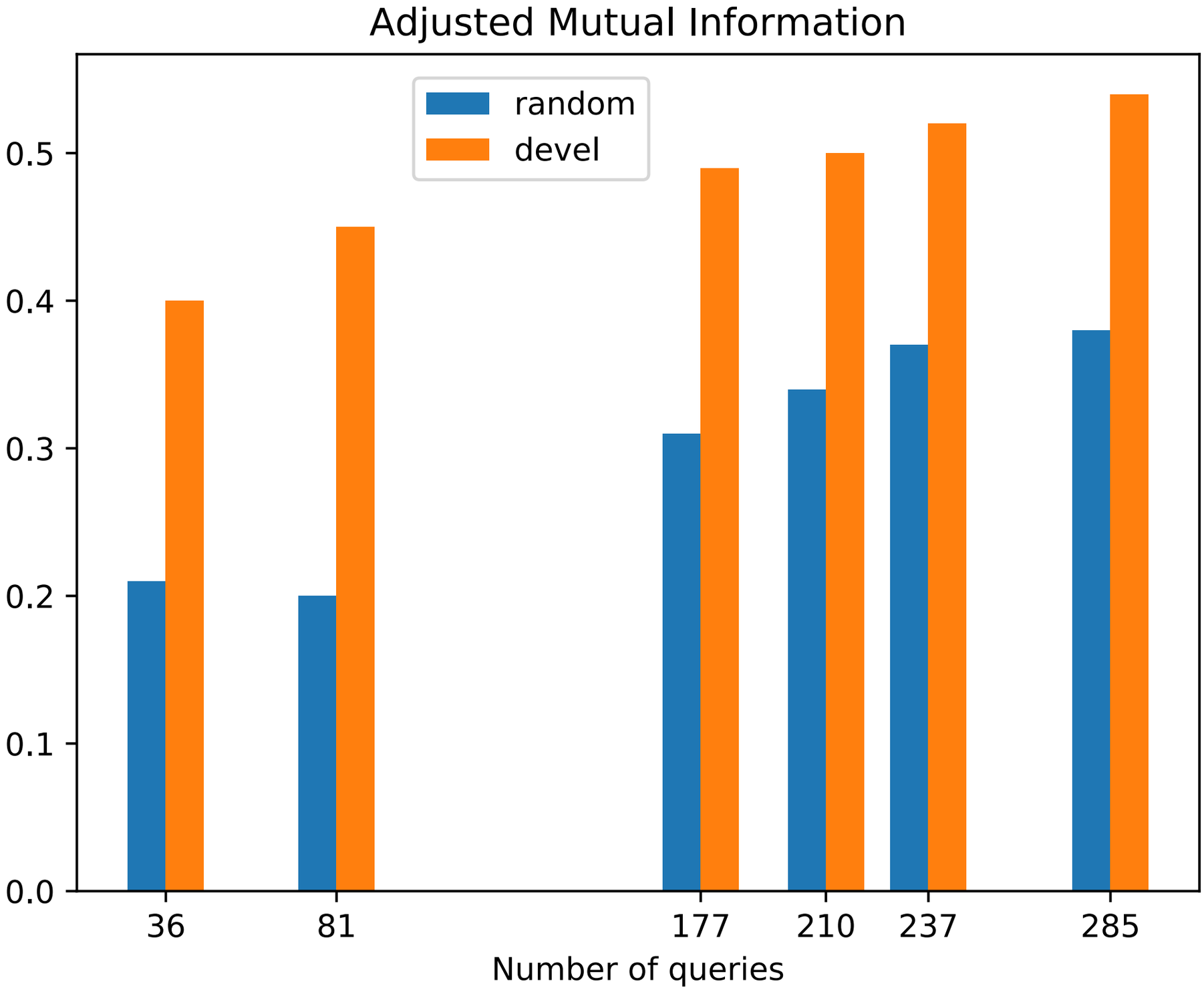




## Developmental



# Experiments: unsupervised recognition



# Credits



Fausto Giunchiglia
(DISI - UNITN)



Stefano Teso
(DISI - UNITN)



Andrea Bontempelli
(DISI - UNITN)

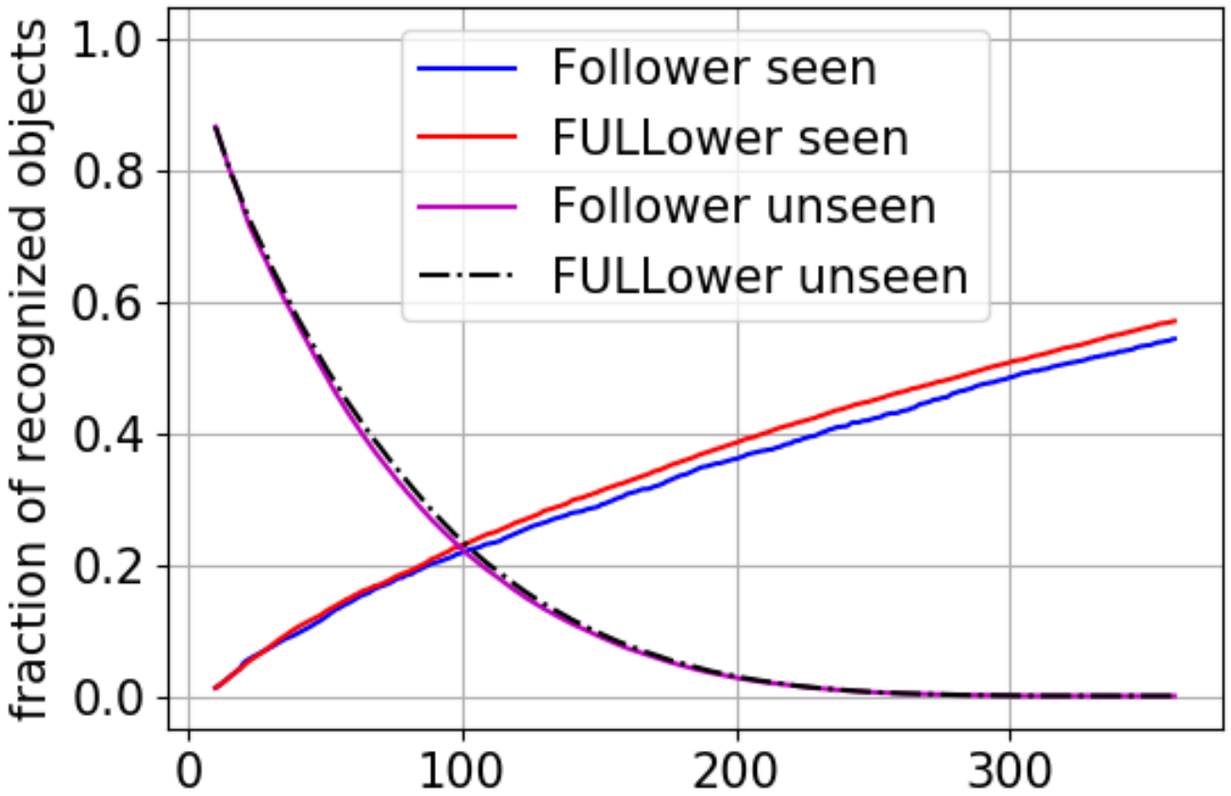


Luca Erculiani
(DISI - UNITN)

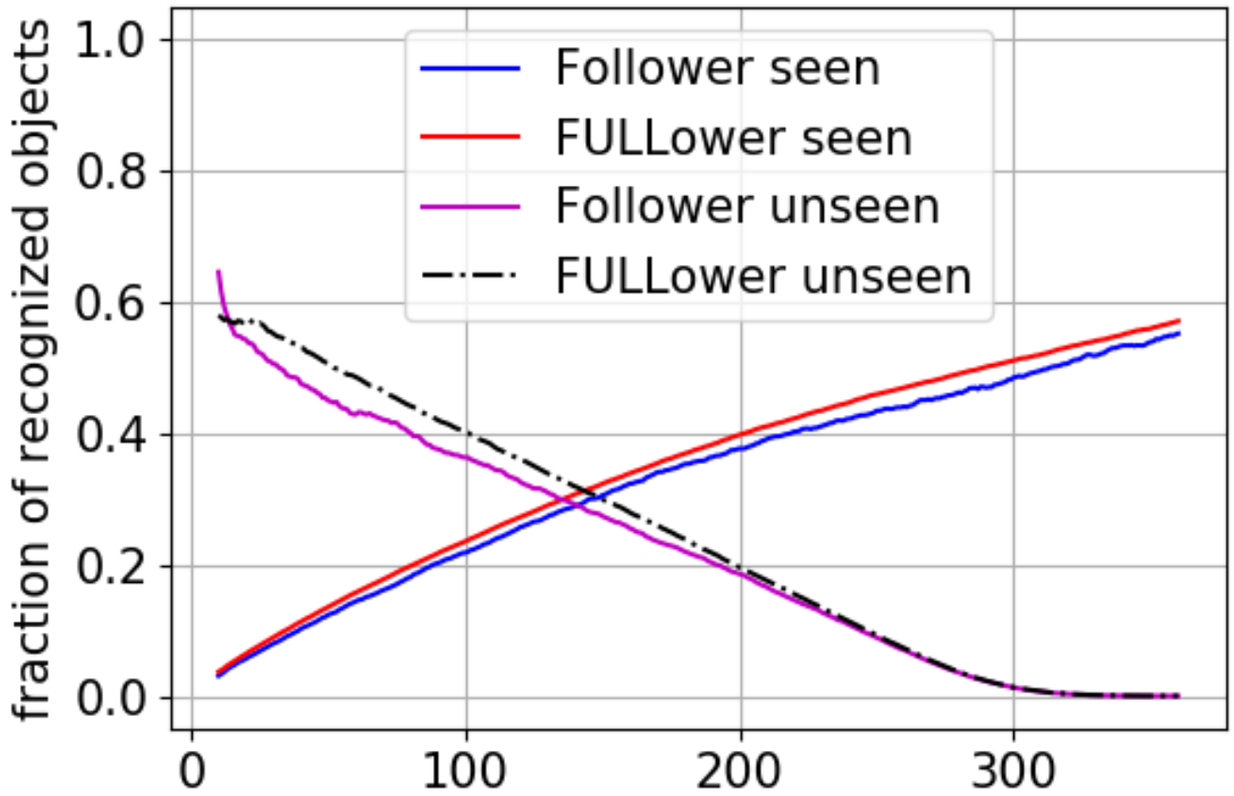
Thank you, questions?

# Experiments: test set accuracy

## Random




## Developmental



# Skeptical learning: formulas

active learning

$$\mathbb{P}(f_{\hat{y}_t}(x_t) - f_{\tilde{y}_t}(x_t) \geq 0) = \Phi\left(\delta_t(x_t)/\sigma_t(x_t)\right)$$

skepticality

$$\mathbb{P}_{\hat{y}_t}(f(x_t) \leq 0 \,|\, x_t) = 1 - \Phi\left(\mu_{\hat{y}_t,t}(x_t)/\sigma_t(x_t)\right)$$