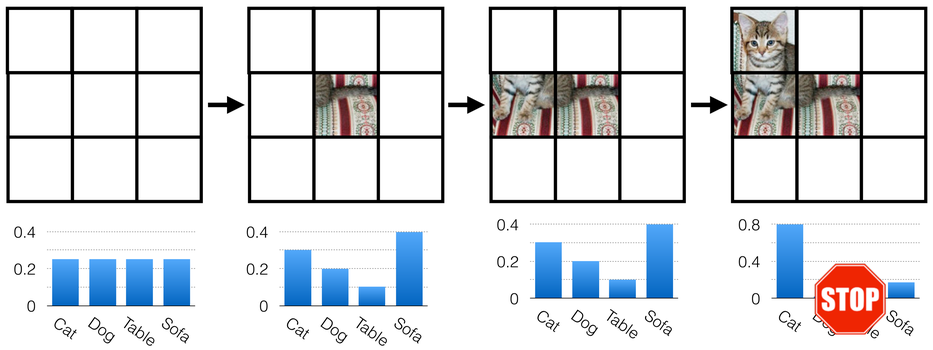
Active Information Acquisition

He He
Paul Mineiro & Nikos Karampatziakis

University of Maryland, College Park
Microsoft Cloud and Information Services Lab

Overview

Dynamically seek information needed most



- **Adaptive**: selection of next information depends on past information and intermediate predictions
- **Cost-efficient**: stop and output results as soon as enough information has been acquired

Problem formulation

- **State:** information acquired so far and intermediate predictions
- **Action**: get a new piece of information or stop (and output current prediction)
- **Loss**: *task loss* + λ · *information cost*

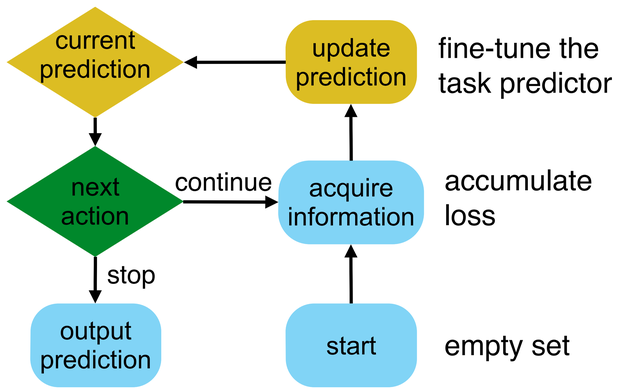
Method

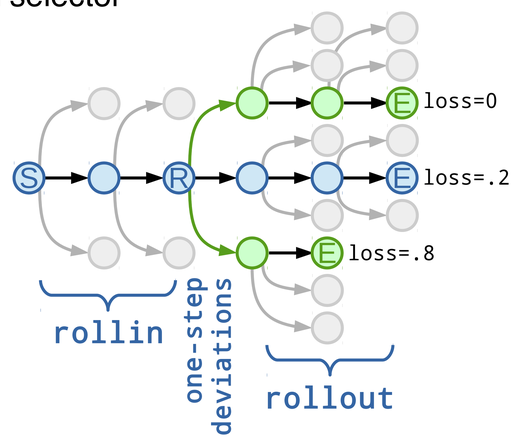
Our goal is to learn

- **Task predictor**: takes in partial input, outputs (intermediate) prediction
- **Information selector**: takes in a state representation, outputs the next action

Learning to Search (Daumé III et al, 2014)

- An imitation learning framework via online cost-sensitive classification
- Explore by rolling in with learned policy; assign credit by rolling out with the reference policy
- **Reference policy:** greedily choose the next action that yields the minimum immediate loss
- Jointly learn the task predictor and the information selector

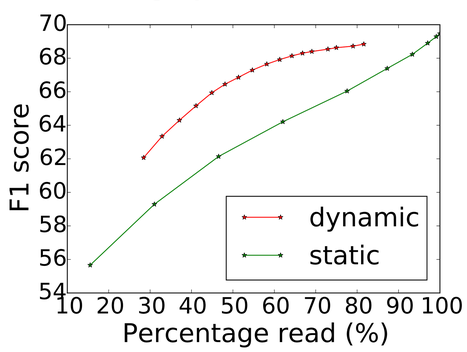




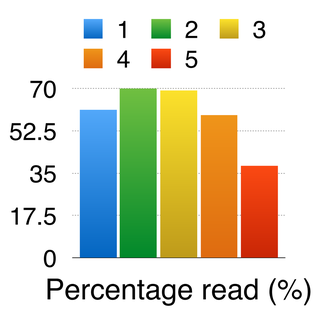
TL;DR

- **When to stop**: sentiment classification on Amazon book reviews
- Read a review from the beginning; 2 actions (stop and continue)
- Task predictor: bag-of-words; one-against-all (5 classes)

Being dynamic is better

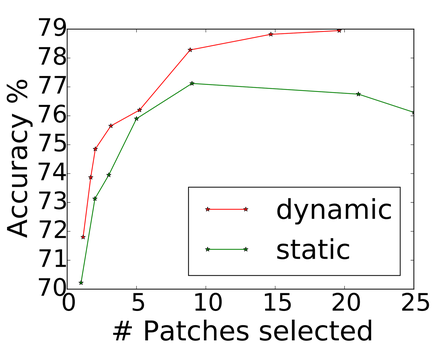


Spend more on hard problems!



TB;DL

- **Where to focus**: image recognition on PASCAL VOC 2011
- Divide an image into 5x5 patches; reveal one patch at a time; 26 actions (patch ID and stop)
- Patch aggregation: linear logistic regression using patch features from last layer of CNN
- Baseline: heuristically selected patches (going from middle to outer part)



Low budget: focus on the middle part; less dynamic
High budget: explore outer part; more dynamic