Near-optimal **Batch Mode** Active Learning and Adaptive Submodular Optimization

Yuxin Chen   Andreas Krause

*Department of Computer Science, ETH Zurich*
Batch Mode Active Learning
Multi-stage Influence Maximization in Social Networks
How should we construct the batches?
THE **BatchGreedy** ALGORITHM

Conditional marginal benefit of an item $s$:

$$\Delta f(s \mid A, y_B) = \mathbb{E}_{y_V} \left[ f(y_{\{s\} \cup A \cup B}) - f(y_{A \cup B}) \mid y_B \right].$$

- **Expectation** over all realizations within the batch
- **Conditioning** on previous observations
**THE BatchGreedy ALGORITHM**

**BatchGreedy** will greedily select the i-th element in the j-th batch

\[
s_{i,j} = \arg \max_{s \in \mathcal{V}} \Delta_f(s \mid \{s_{1,j}, \ldots, s_{i-1,j}\}, y_{\mathcal{B}}) \]

\[
\text{the } j_{th} \text{ batch } \mathcal{A}
\]
**BatchGreedy** vs. Optimal Batch

Cost of **BatchGreedy**

Cost of optimal **batch policy**
**BatchGreedy** vs. Optimal Batch

Cost of **BatchGreedy** \( \leq O(\ln Q) \). Cost of optimal **batch policy**
How many extra items will we select?
BatchGreedy VS. SEQUENTIAL

Cost of BatchGreedy

Cost of optimal sequential policy
BatchGreedy VS. SEQUENTIAL

Cost of BatchGreedy

Cost of optimal sequential policy

COMPETITIVE
EXPERIMENTAL RESULTS
EXPERIMENTAL RESULTS

Number of labels requested vs. % Mistakes
EXPERIMENTAL RESULTS

Number of labels requested vs % Mistakes

- Random strategy
- Graph of data points showing a decreasing trend

Diagram showing a scatter plot with a line of best fit and a legend indicating 'random'.
EXPERIMENTAL RESULTS

Number of labels requested vs. % Mistakes

- Random
- Sequential
EXPERIMENTAL RESULTS
EXPERIMENTAL RESULTS

- Number of labels requested
- % Mistakes

Comparing methods:
- 10–batch greedy
- Sequential
- KLR–BMAL
- Random
EXPERIMENTAL RESULTS

- Graph showing data points and labels.
- Another line graph showing % item not covered vs Number of items selected.
EXPERIMENTAL RESULTS

- % Mistakes vs. Number of labels requested
  - KLR-BMAL
  - random
  - sequential
  - 10-batch greedy

- % item not covered vs. Number of items selected
  - Non-adaptive
EXPERIMENTAL RESULTS
EXPERIMENTAL RESULTS

![Graphical representation of experimental results]

- **Number of labels requested**
  - Batch greedy
  - Random
  - KLR-BMAL

- **% Mistakes**
  - Sequential
  - 10-batch greedy

- **% item not covered**
  - Sequential
  - 10-batch
  - Non-adaptive
EXPERIMENTAL RESULTS

% Mistakes

Number of labels requested

% item not covered

Number of items selected

sequential

10–batch greedy

KLR–BMAL

random

100–batch

Non–adaptive

sequential

10–batch

Non–adaptive
Code Available

www.inf.ethz.ch/~chenyux/icml13/bmal-src.zip

Thanks for your attention

Come to see our poster!