Efficient Nonmyopic Active Search

Jiang, Malkomes, Converse, Shofner, Moseley and Garnett

STA 4273/CSC 2547 Paper Presentation

Presented by: Zain Hasan & Daniel Hidru

Active Search

- Sequentially locating as many members of a particular class as possible - *targets that belong to a rare class*

$$\mathcal{D} \triangleq \{(x_i, y_i)\} \qquad\qquad y \triangleq \mathbb{1}\{x \in \mathcal{R}\}$$

- Active search is Bayesian optimization with binary rewards and cumulative regret (budget).

Analogy for Active Search

- Writing a Literature Review is an active search process
  - Limited amount of papers you can read (budget)
  - Reading papers you know are relevant (exploitation)
  - Reading papers that might be relevant in the hope that you find more relevant papers (exploration)

Budget (Cumulative regret)

- You have limited time (deadline) and resources.
- Have to balance between exploration and exploitation to maximize utility for binary y = {0,1}:

$$\boxed{u(\mathcal{D}) \triangleq \sum_{y_i \in \mathcal{D}} y_i,}$$

- Which counts the number of targets in chosen set (ie. Relevant papers included in review)
- Want to determine/approximate some optimal policy of picking points that maximizes utility

Myopic vs. Nonmyopic

- Myopic search: consider the effect of only potential immediate choices
  - Easier, lower runtime complexity, but short-sighted

- Nonmyopic search: consider impact of all selected points, immediate and future
  - Harder, more complex, but potentially better results

Contributions of Paper

1. Prove that active search, that approximates the optimal policy, is hard to do by finding its runtime complexity via a proof

2. Suggest an efficient nonmyopic search algorithm

Background for Algorithm

- Optimal Bayesian Decision/Policy:
  - Posterior prob. of a point belonging to desired y = 1 class

$$\boxed{\Pr(y = 1 \mid x, \mathcal{D})}$$

- Choose next points maximizing the expected number of targets found at termination, given i - 1 previous observations:

$$\boxed{x_i^* = \underset{x_i \in \mathcal{X} \setminus \mathcal{D}_{i-1}}{\arg\max} \; \mathbb{E}\big[u(\mathcal{D}_t) \mid x_i, \mathcal{D}_{i-1}\big]}$$

Expected utility: 1 query left

$$\boxed{\begin{aligned}\mathbb{E}\left[u(\mathcal{D}_t) \mid x_t, \mathcal{D}_{t-1}\right] &= \textstyle\sum_{y_t} u(\mathcal{D}_t)\Pr(y_t \mid x_t, \mathcal{D}_{t-1}) \\ &= u(\mathcal{D}_{t-1}) + \Pr(y_t = 1 \mid x_t, \mathcal{D}_{t-1}). \qquad (3)\end{aligned}}$$

Expected utility of selecting x_t, given previous selections (D_{t-1}) =

Reward for previous selections + Expected reward of current selection

- Pure exploitation because there are no more queries to make

Expected utility: 2 queries left

$$\begin{aligned}\mathbb{E}\big[u(\mathcal{D}_t) \mid x_{t-1}, \mathcal{D}_{t-2}\big] = u(\mathcal{D}_{t-2}) &+ \\ \Pr(y_{t-1} = 1 \mid x_{t-1}, \mathcal{D}_{t-2}) &+ \\ \mathbb{E}_{y_{t-1}}\big[\max_{x_t} \Pr(y_t = 1 \mid x_t, \mathcal{D}_{t-1})\big]&.\end{aligned}$$

Expected utility of selecting x_{t-1}, given previous selections (D_{t-2}) =

Reward for previous selections + Expected reward of current selection + Expected reward for final selection given outcome of current selection

- Natural trade off between exploitation (2nd term) and exploration (3rd term)

Expected utility: t-i+1 queries left

$$\mathbb{E}\big[u(\mathcal{D}_t) \mid x_i, \mathcal{D}_{i-1}\big] = u(\mathcal{D}_{i-1}) + \underbrace{\Pr(y_i = 1 \mid x_i, \mathcal{D}_{i-1})}_{\text{exploitation, } < 1} + \underbrace{\mathbb{E}_{y_i}\Big[\max_{x'} \mathbb{E}\big[u(\mathcal{D}_t \setminus \mathcal{D}_i) \mid x', \mathcal{D}_i\big]\Big]}_{\text{exploration, } < t-i}$$

Expected utility of selecting x_i, given previous selections (D_{i-1}) =

Reward for previous selections + Expected reward of current selection + Expected reward for <u>remaining</u> selections given outcome of current selection

- Can compute this expectation recursively
  - Cost: exponential in the number of future queries - O((2n)^l)

Hardness of Approximation

Theorem 1. *There is no polynomial-time active search policy with a constant factor approximation ratio for optimizing the expected utility.*

Efficient Nonmyopic Search (ENS): t-i+1 queries left

$$\max_{x'} \mathbb{E}\big[u(\mathcal{D}_t \backslash \mathcal{D}_i) \mid x', \mathcal{D}_i\big] \approx \textstyle\sum'_{t-i} \Pr(y = 1 \mid x, \mathcal{D}_i),$$

$$\begin{aligned}\mathbb{E}\big[u(\mathcal{D}_t) \mid x_i, \mathcal{D}_{i-1}\big] \approx u(\mathcal{D}_{i-1}) &+ \\ \Pr(y_i = 1 \mid x_i, \mathcal{D}_{i-1}) &+ \\ \underbrace{\mathbb{E}_{y_i}\Big[\textstyle\sum'_{t-i} \Pr\big(y = 1 \mid x, \mathcal{D}_i\big)\Big]}_{\text{exploration}, < t-i}\end{aligned}$$

Expected utility of selecting x_i, given previous selections (D_{i-1}) ≈

Reward for previous selections + Expected reward of current selection + Expected reward for remaining selections given they are <u>selected as a batch</u>

- Assumption: the labels of all unlabeled points are conditionally independent
  - Needed to reduce the final term to a sum of marginal probabilities

Assumptions for efficiency improvements

1. Updating the model only affects a limited number of samples.
2. Observing a new negative point will not raise the probability of any other point being a target.
3. Able to bound the maximum probability of the unlabeled data conditioned on the future selection of additional targets.

Representative experiment: CiteSeer data

- Data:
  - 39,788 computer science papers published in the top 50 venues
  - 2,190 (5.5%) are NIPS publications
- Goal: Find the most NIPS publications given a budget t=500
- Model: k-NN with k=50
  - Easy to update
  - Consistent with efficiency assumptions
- Features: graph PCA on the citation network using the first 20 principal components

Results: All 500 queries

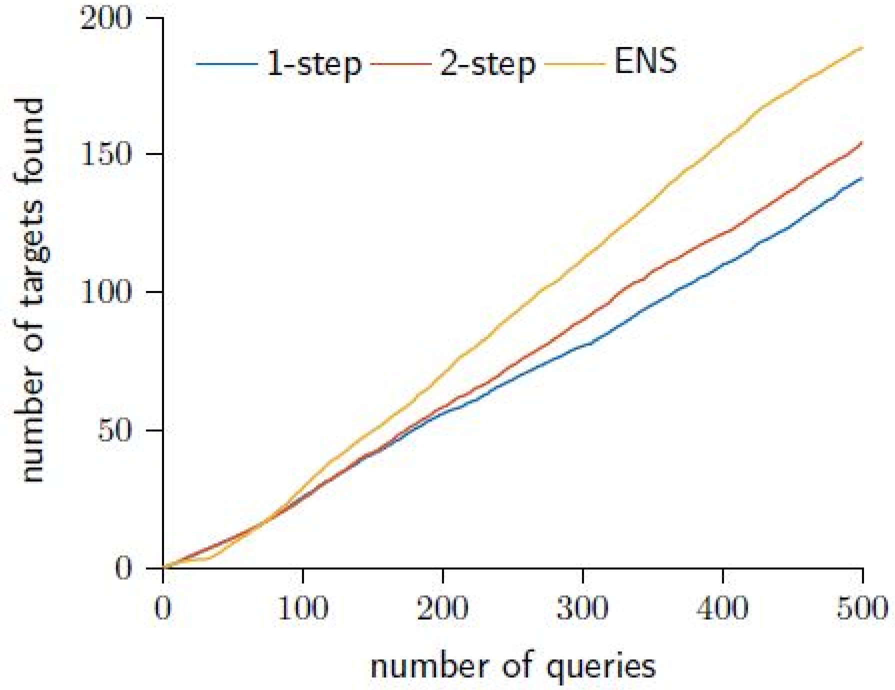

Image: https://bayesopt.github.io/slides/2016/ContributedGarnett.pdf

Results: First 80 queries

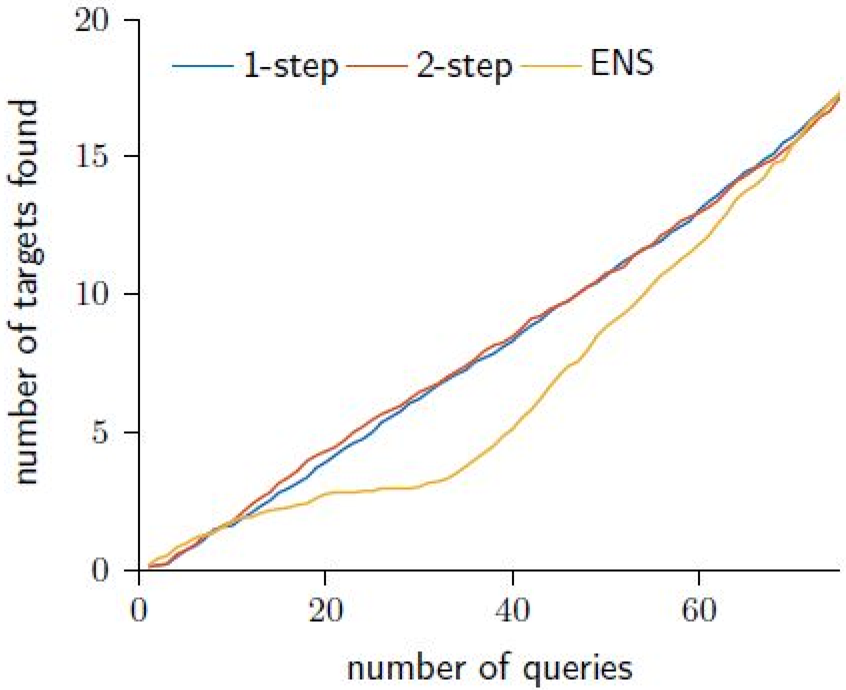

Image: https://bayesopt.github.io/slides/2016/ContributedGarnett.pdf

Results: Different budgets

	query number				
policy	100	300	500	700	900
one-step	25.5	80.5	141	209	273
two-step	24.9	89.8	155	220	287
ENS–900	25.9	94.3	163	239	**308**
ENS–700	28.0	105	188	**259**	
ENS–500	28.7	**112**	**189**		
ENS–300	26.4	105			
ENS–100	**30.7**				

Image: https://bayesopt.github.io/slides/2016/ContributedGarnett.pdf

Relationship to other fields of research

- Active learning: train a high performing model with a few selected examples
  - AS: find elements of a rare class with a few selected choices
- Multi-armed bandit: maximize expected score given limited resources
  - AS: items are correlated and can only be selected once
  - ENS similar to knowledge gradient policy (Frazier et al., 2008)
- Bayesian optimization: global optimization using sequential choices
  - AS: special case with binary observations and cumulative reward
  - ENS similar to GLASSES algorithm (González et al., 2016)

Limitations (related to this course) and future work

- Active Search/ENS Approach
  - Can’t select the same element multiple times
    - Difficult to apply to reinforcement learning where the same action can be repeated
  - Can’t work in a continuous object domain
    - Needs discrete objects that can’t be selected multiple times to avoid selecting objects that are arbitrarily close to a previously selected item
  - True reward does not depend on previous actions
    - The order of the decisions affects your performance in reinforcement learning
- Bayesian Optimization
  - Probability models need to be updated multiple times before each selection
    - Costly to retrain neural networks (idea: update with a few gradient steps)
  - Difficult to work with continuous labels/rewards
    - Challenging to integrate the expected future reward (idea: estimate expectation)

Summary

- Efficient Nonmyopic Search outperforms myopic search in the active search problem by considering the benefit of exploration associated with the rewards of future queries.
- The key idea will be difficult to utilize in our course projects because it depends on many of the constraints imposed by the problem definition.

References

- Jiang, S., Malkomes, G., Converse, G., Shofner, A., Moseley, B. and Garnett, R., 2017, July. Efficient Nonmyopic Active Search. In International Conference on Machine Learning (pp. 1714-1723).
- Garnett, R., 2016, October. Efficient Nonmyopic Active Search. https://bayesopt.github.io/slides/2016/ContributedGarnett.pdf
- Frazier, P.I., Powell, W.B. and Dayanik, S., 2008. A knowledge-gradient policy for sequential information collection. SIAM Journal on Control and Optimization, 47(5), pp.2410-2439.
- González, J., Osborne, M. and Lawrence, N., 2016, May. GLASSES: Relieving the myopia of Bayesian optimisation. In Artificial Intelligence and Statistics (pp. 790-799).