Proceedings of Machine Learning Research vol 247:1–2, 2024 37th Annual Conference on Learning Theory

Mode Estimation with Partial Feedback

Charles Arnal* CHARLES.ARNAL@INRIA.FR
Datashape, Inria
Vivien Cabannes* VIVC@META.COM
FAIR, Meta AI
Vianney Perchet CVIANNEY.PERCHET@ENSAE.EDU
Crest, ENSAE

Editors: Shipra Agrawal and Aaron Roth

Abstract

The combination of lightly supervised pre-training and online fine-tuning has played a key role in recent AI developments. These new learning pipelines call for new theoretical frameworks. In this paper, we formalize core aspects of weakly supervised and active learning with a simple problem: the estimation of the mode of a distribution using partial feedback. Variations of this problem naturally occur in many instances where exhaustive information on a dataset is either too expensive or downright impossible to obtain, e.g. due to the limitations of human annotators, but partial, hierarchical information is available (Ziegler et al., 2020; Zhu et al., 2023; Cabannes et al., 2022; Cesa-Bianchi et al., 2019; Jamieson and Nowak, 2011; Braverman et al., 2019; Cesa-Bianchi et al., 2006; Gangaputra and Geman, 2006; Fiez et al., 2019).

More formally, the practitioner has access to samples (Y_j) drawn iid from a discrete distribution $p \in \mathbb{P}(\mathcal{Y})$, where $\mathcal{Y} = \{y_1, \dots, y_m\}$ and $p(y_i) \geq p(y_{i+1})$, but cannot directly observe the realizations Y_j; instead, at each time $t \in \mathbb{N}$, they select an index j_t and a set $S_t \subset \mathcal{Y}$ and observe the value of the *query* $\mathbf{1}_{Y_{j_t} \in S_t}$. The task is to accurately estimate the mode $y_1 = \arg\max_{y \in \mathcal{Y}} p(y)$ of the distribution p using as few queries as possible.

A naive solution is to simply identify each sample Y_j exactly through a binary search consuming $\lceil \log_2(m) \rceil$ queries, and to output the most frequent class among the identified samples as an estimation of the true mode at any time t. The asymptotic number T of queries needed by this algorithm to identify the mode with probability of error at most δ, as $\delta \to 0$, is

$$T \leq (\Delta_2^{-2} + \Delta_2^{-2} \log_2(m)) \log(1/\delta),$$

where $\Delta_i^2 := -\ln(1 - (\sqrt{p(y_1)} - \sqrt{p(y_i)})^2)$ and we ignore multiplicative constants. We improve upon this baseline by combining three ideas.

- We identify samples using a learnt, asymptotically optimal entropy-based dichotomic search tree (Huffman, 1952; Vitter, 1987). This allows us to replace $\log_2(m)$ by $-\sum p(y_i) \log_2(p(y_i))$.
- We do not fully identify each sample; instead, we stop our search at a depth in the tree roughly equal to that of the mode. Nodes below this depth correspond to sets S of classes whose estimated probability is smaller than that of the mode. This allows us to replace $\log_2(p(y_i))$ by $\log_2(p(y_1))$.
- We use bandits-inspired techniques to discard classes $y \in \mathcal{Y}$ as soon as they seem unlikely to be the mode (Even-Dar et al., 2006; Bubeck et al., 2010). This allows us to replace Δ_2^{-1} by Δ_i^{-1}.

The asymptotic expected number of queries required by the resulting algorithm to identify the mode with probability of error at most δ satisfies

$$E[T] \leq (\Delta_2^{-2} + (\sum_{i \geq 1} p(y_i) \Delta_i^{-2}) |\log_2 p(y_1)|) \log(1/\delta),$$

once again ignoring multiplicative constants. All our algorithms and some code to run experiments are available online at https://github.com/VivienCabannes/mepf.

Keywords: Active Learning, Partial Feedback, Coding, Coarse Search, Best Arm Identification.

 *These authors contributed equally to this work. .

Acknowledgments

The authors would like to thank Gilles Blanchard, Remi Jezequiel, Marc Jourdan, Nadav Merlis, and Karen Ullrich for fruitful discussions.

References

Mark Braverman, Jieming Mao, and Yuval Peres. Sorted top-k in rounds. In *COLT*, 2019.

Sébastien Bubeck, Rémi Munos, and Gilles Stoltz. Pure exploration for multi-armed bandit problems. In *COLT*, 2010.

Vivien Cabannes, Francis Bach, Vianney Perchet, and Alessandro Rudi. Active labeling: Streaming stochastic gradients. In *NeurIPS*, 2022.

Nicolò Cesa-Bianchi, Claudio Gentile, and Luca Zaniboni. Incremental algorithms for hierarchical classification. *Journal of Machine Learning Research*, 2006.

Nicolò Cesa-Bianchi, Tommaso Cesari, and Vianney Perchet. Dynamic pricing with finitely many unknown valuations. In *ALT*, 2019.

Eyal Even-Dar, Shie Mannor, and Yishay Mansour. Action elimination and stopping conditions for the multi-armed bandit and reinforcement learning problems. *Journal of Machine Learning Research*, 2006.

Tanner Fiez, Lalit Jain, Kevin Jamieson, and Lillian Ratliff. Sequential experimental design for transductive linear bandits. In *NeurIPS*, 2019.

Sachin Gangaputra and Donald Geman. A design principle for coarse-to-fine classification. In *Conference on Computer Vision and Pattern Recognition*, 2006.

David Huffman. A method for the construction of minimum-redundancy codes. *Proceedings of the Institute of Radio Engineers*, 40(9):1098–1101, 1952.

Kevin Jamieson and Robert Nowak. Active ranking using pairwise comparisons. In *NeurIPS*, 2011.

Jeffrey Vitter. Design and analysis of dynamic Huffman codes. *Journal of the ACM*, 34(4):825–845, 1987.

Banghua Zhu, Michael Jordan, and Jiantao Jiao. Principled reinforcement learning with human feedback from pairwise or k-wise comparisons. In *ICML*, 2023.

Daniel Ziegler, Nisan Stiennon, Jeffrey Wu, Tom Brown, Alec Radford, Dario Amodei, Paul Christiano, and Geoffrey Irving. Fine-tuning language models from human preferences. Technical report, OpenAI, 2020.