



Efficient Optimal Learning for Contextual Bandits

Miroslav Dudik, Daniel Hsu, Satyen Kale, Nikos Karampatziakis, John Langford, Lev Reyzin, Tong Zhang

(Submitted on 13 Jun 2011)

We address the problem of learning in an online setting where the learner repeatedly observes features, selects among a set of actions, and receives reward for the action taken. We provide the first efficient algorithm with an optimal regret. Our algorithm uses a cost sensitive classification learner as an oracle and has a running time $\mathrm{polylog}(N)$, where N is the number of classification rules among which the oracle might choose. This is exponentially faster than all previous algorithms that achieve optimal regret in this setting. Our formulation also enables us to create an algorithm with regret that is additive rather than multiplicative in feedback delay as in all previous work.

Subjects: **Learning (cs.LG)**; Artificial Intelligence (cs.AI); Machine Learning (stat.ML)

Cite as: [arXiv:1106.2369](#) [cs.LG]
(or [arXiv:1106.2369v1](#) [cs.LG] for this version)

Submission history

From: Daniel Hsu [[view email](#)]
[v1] Mon, 13 Jun 2011 01:57:52 GMT (51kb)

[Which authors of this paper are endorsers?](#)

Link back to: [arXiv](#), [form interface](#), [contact](#).

Download:

- [PDF](#)
- [PostScript](#)
- [Other formats](#)

Current browse context:

cs.LG

[< prev](#) | [next >](#)

[new](#) | [recent](#) | [1106](#)

Change to browse by:

cs

[cs.AI](#)

stat

[stat.ML](#)

References & Citations

- [NASA ADS](#)

DBLP - CS Bibliography

[listing](#) | [bibtex](#)

[Miroslav Dudík](#)

[Daniel Hsu](#)

[Satyen Kale](#)

[Nikos Karampatziakis](#)

[John Langford](#)

...

Bookmark (what is this?)

