



Using More Data to Speed-up Training Time

Shai Shalev-Shwartz, Ohad Shamir, Eran Tromer

(Submitted on 6 Jun 2011 (v1), last revised 15 Jun 2011 (this version, v2))

In many recent applications, data is plentiful. By now, we have a rather clear understanding of how more data can be used to improve the accuracy of learning algorithms. Recently, there has been a growing interest in understanding how more data can be leveraged to reduce the required training runtime. In this paper, we study the runtime of learning as a function of the number of available training examples, and underscore the main high-level techniques. We provide some initial positive results showing that the runtime can decrease exponentially while only requiring a polynomial growth of the number of examples, and spell-out several interesting open problems.

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

Cite as: **arXiv:1106.1216 [cs.LG]**

(or **arXiv:1106.1216v2 [cs.LG]** for this version)

Submission history

From: Ohad Shamir [[view email](#)]

[v1] Mon, 6 Jun 2011 23:55:00 GMT (19kb)

[v2] Wed, 15 Jun 2011 01:17:56 GMT (18kb)

[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

stat

[stat.ML](#)

References & Citations

- [NASA ADS](#)

DBLP - CS Bibliography

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

Shai Shalev-Shwartz

Ohad Shamir

Eran Tromer

Bookmark (what is this?)

