

The MIT Press

Journals

[Sign In / Register](#)

[Books](#)

[Journals](#)

[Digital](#)

[Resources](#)

[About](#)



[Home](#) | [Computational Linguistics](#) | [List Article navigation of Issues](#) | [Volume 38 , No. 3](#) | [Empirical Risk Minimization for Probabilistic Grammars: Sample Complexity and Hardness of Learning](#)



Quarterly (March, June, September, December)

160pp. per issue

6 3/4 x 10

Founded: 1974

2018 Impact Factor: 1.319

2018 Google Scholar h5-index: 32

ISSN: 0891-2017

E-ISSN: 1530-9312

Journal Resources

- [Editorial Info](#)
- [Abstracting and Indexing](#)
- [Release Schedule](#)
- [Advertising Info](#)

Author Resources

- [Submission Guidelines](#)

Empirical Risk Minimization for Probabilistic Grammars: Sample Complexity and Hardness of Learning

[Shay B. Cohen](#) and [Noah A. Smith](#)

Posted Online December 08, 2011
https://doi.org/10.1162/COLI_a_00092

© 2012 Association for Computational Linguistics

Computational Linguistics
Volume 38 | Issue 3 | September 2012
p.479-526

 [Download Options](#) >

[Abstract](#) [Full Text](#) [Authors](#)

Probabilistic grammars are generative statistical models that are useful for compositional and

Publication
 Agreement
 Author Reprints

Reader Resources

Rights and Permissions
 Most Read
 Most Cited

More About Computational Linguistics ▼

Metrics ▼



3 Total

citations

0 Recent

citations

0.77 Field Citation

Ratio

n/a Relative

Citation Ratio

Open Access ▼



Computational Linguistics Computational Linguistics is Open Access. All content is freely available in electronic format (Full text HTML, PDF, and PDF Plus) to

sequential structures. They are used ubiquitously in computational linguistics. We present a framework, reminiscent of structural risk minimization, for empirical risk minimization of probabilistic grammars using the log-loss. We derive sample complexity bounds in this framework that apply both to the supervised setting and the unsupervised setting. By making assumptions about the underlying distribution that are appropriate for natural language scenarios, we are able to derive distribution-dependent sample complexity bounds for probabilistic grammars. We also give simple algorithms for carrying out empirical risk minimization using this framework in both the supervised and unsupervised settings. In the unsupervised case, we show that the problem of minimizing empirical risk is NP-hard. We therefore suggest an approximate algorithm, similar to expectation-maximization, to minimize the empirical risk.

Forthcoming

Most Read

[See More](#)

Lexicon-Based Methods for Sentiment Analysis (14019 times)
 Maite Taboada et al.
 Computational Linguistics
 Volume: 37, Issue: 2, pp. 267-307

Computational Linguistics and Deep Learning (10513 times)
 Christopher D. Manning
 Computational Linguistics
 Volume: 41, Issue: 4, pp. 701-707


Near-Synonymy and Lexical Choice (3658 times)
 Philip Edmonds et al.
 Computational Linguistics
 Volume: 28, Issue: 2, pp. 105-144


(Note that the Most Read numbers are based on the number of full text downloads over the last 12 months.)


Most Cited

[See More](#)

readers across the globe. All articles are published under a [CC BY-NC-ND 4.0 license](#). For more information on allowed uses, please view the [CC license](#). [Support OA at MITP](#)



 **Lexicon-Based Methods for Sentiment Analysis** (436 times)
Maite Taboada et al.
Computational Linguistics
Volume: 37, Issue: 2, pp. 267-307

 **A Systematic Comparison of Various Statistical Alignment Models** (174 times)
Franz Josef Och et al.
Computational Linguistics
Volume: 29, Issue: 1, pp. 19-51

 **Opinion Word Expansion and Target Extraction through Double Propagation** (147 times)
Guang Qiu et al.
Computational Linguistics
Volume: 37, Issue: 1, pp. 9-27

(Note that the Most Cited numbers are based on Crossref's [Cited-by service](#) and reflect citation information for the past 24 months.)

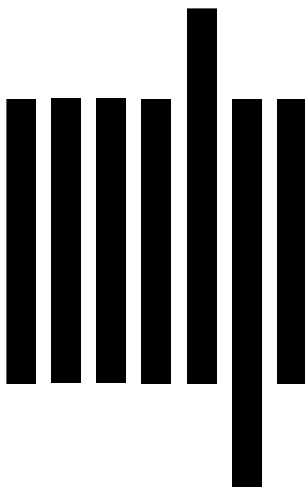
 **Download Options** >

Favorite  Sign up for Alerts 

Download Citation  RSS TOC 

RSS Citation  Submit your article

[Support OA at MITP](#) 



Journals

Terms & Conditions

Privacy Statement

Contact Us

Books

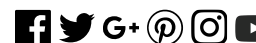
US

UK

Connect

One Rogers Street
Cambridge MA
02142-1209

Suite 2, 1 Duchess
Street London,
W1W 6AN, UK



© 2018 The MIT Press
Technology Partner:
[Atypon Systems, Inc.](#)
[CrossRef Member](#)
[COUNTER Member](#)
The MIT Press colophon is registered in the

U.S. Patent and Trademark Office

