Improving Distributional Similarity with Lessons Learned from Word Embeddings

Omer Levy, Yoav Goldberg, Ido Dagan

Abstract

Recent trends suggest that neural-network-inspired word embedding models outperform traditional count-based distributional models on word similarity and analogy detection tasks. We reveal that much of the performance gains of word embeddings are due to certain system design choices and hyperparameter optimizations, rather than the embedding algorithms themselves. Furthermore, we show that these modifications can be transferred to traditional distributional models, yielding similar gains. In contrast to prior reports, we observe mostly local or insignificant performance differences between the methods, with no global advantage to any single approach over the others.

Anthology ID:
Q15-1016
Volume:
Transactions of the Association for Computational Linguistics, Volume 3
Month:
Year:
2015
Address:
Cambridge, MA
Venue:
TACL
SIG:
Publisher:
MIT Press
Note:
Pages:
211–225
Language:
URL:
https://aclanthology.org/Q15-1016
DOI:
10.1162/tacl_a_00134
Bibkey:
levy-etal-2015-improving
Cite (ACL):
Omer Levy, Yoav Goldberg, and Ido Dagan. 2015. Improving Distributional Similarity with Lessons Learned from Word Embeddings. *Transactions of the Association for Computational Linguistics*, 3:211–225.
Cite (Informal):
Improving Distributional Similarity with Lessons Learned from Word Embeddings (Levy et al., TACL 2015)
Copy Citation:
BibTeX Markdown MODS XML Endnote More options…
PDF:
https://aclanthology.org/Q15-1016.pdf

PDF Cite Search

Export citation

×

- BibTeX
- MODS XML
- Endnote
- Preformatted

```
@article{levy-etal-2015-improving,
    title = "Improving Distributional Similarity with Lessons Learned from Word Embeddings",
    author = "Levy, Omer  and
      Goldberg, Yoav  and
      Dagan, Ido",
    journal = "Transactions of the Association for Computational Linguistics",
    volume = "3",
    year = "2015",
    address = "Cambridge, MA",
    publisher = "MIT Press",
    url = "https://aclanthology.org/Q15-1016",
    doi = "10.1162/tacl_a_00134",
    pages = "211--225",
    abstract = "Recent trends suggest that neural-network-inspired word embedding models outperform traditional count-based distributiona
}
```

Download as File Copy to Clipboard

```
<?xml version="1.0" encoding="UTF-8"?>
<modsCollection xmlns="http://www.loc.gov/mods/v3">
<mods ID="levy-etal-2015-improving">
    <titleInfo>
        <title>Improving Distributional Similarity with Lessons Learned from Word Embeddings</title>
    </titleInfo>
    <name type="personal">
        <namePart type="given">Omer</namePart>
        <namePart type="family">Levy</namePart>
        <role>
            <roleTerm authority="marcrelator" type="text">author</roleTerm>
        </role>
    </name>
    <name type="personal">
        <namePart type="given">Yoav</namePart>
        <namePart type="family">Goldberg</namePart>
        <role>
            <roleTerm authority="marcrelator" type="text">author</roleTerm>
        </role>
    </name>
    <name type="personal">
        <namePart type="given">Ido</namePart>
        <namePart type="family">Dagan</namePart>
        <role>
            <roleTerm authority="marcrelator" type="text">author</roleTerm>
        </role>
    </name>
    <originInfo>
        <dateIssued>2015</dateIssued>
    </originInfo>
    <typeOfResource>text</typeOfResource>
    <genre authority="bibutilsgt">journal article</genre>
    <relatedItem type="host">
        <titleInfo>
            <title>Transactions of the Association for Computational Linguistics</title>
        </titleInfo>
        <originInfo>
            <issuance>continuing</issuance>
            <publisher>MIT Press</publisher>
            <place>
                <placeTerm type="text">Cambridge, MA</placeTerm>
            </place>
        </originInfo>
        <genre authority="marcgt">periodical</genre>
        <genre authority="bibutilsgt">academic journal</genre>
    </relatedItem>
    <abstract>Recent trends suggest that neural-network-inspired word embedding models outperform traditional count-based distributional</abstract>
    <identifier type="citekey">levy-etal-2015-improving</identifier>
    <identifier type="doi">10.1162/tacl_a_00134</identifier>
    <location>
        <url>https://aclanthology.org/Q15-1016</url>
    </location>
    <part>
        <date>2015</date>
        <detail type="volume"><number>3</number></detail>
        <extent unit="page">
            <start>211</start>
            <end>225</end>
        </extent>
    </part>
</mods>
</modsCollection>
```


Download as File | Copy to Clipboard


```
%0 Journal Article
%T Improving Distributional Similarity with Lessons Learned from Word Embeddings
%A Levy, Omer
%A Goldberg, Yoav
%A Dagan, Ido
%J Transactions of the Association for Computational Linguistics
%D 2015
%V 3
%I MIT Press
%C Cambridge, MA
%F levy-etal-2015-improving
%X Recent trends suggest that neural-network-inspired word embedding models outperform traditional count-based distributional models on w
%R 10.1162/tacl_a_00134
%U https://aclanthology.org/Q15-1016
%U https://doi.org/10.1162/tacl_a_00134
%P 211-225
```


Download as File | Copy to Clipboard

Markdown (Informal)


```
[Improving Distributional Similarity with Lessons Learned from Word Embeddings](https://aclanthology.org/Q15-1016) (Levy et al., TACL 2015)
```


- Improving Distributional Similarity with Lessons Learned from Word Embeddings (Levy et al., TACL 2015)

ACL

- Omer Levy, Yoav Goldberg, and Ido Dagan. 2015. Improving Distributional Similarity with Lessons Learned from Word Embeddings. *Transactions of the Association for Computational Linguistics*, 3:211–225.

Copy Markdown to Clipboard | Copy ACL to Clipboard