Evaluation methods for unsupervised **word** embeddings

T Schnabel, I Labutov, D Mimno… - Proceedings of the 2015 …, 2015 - aclweb.org

… These methods re- duce **bias**, provide greater insight, and allow us to solicit data-driven relevance judgments … Neural **word** embeddings represent meaning via geometry … provides vector rep- resentations of words such that the relationship be- tween two **vectors** mirrors the …

☆ Cited by 225 Related articles All 11 versions »

Deep contextualized **word** representations

ME Peters, M Neumann, M Iyyer, M Gardner… - arXiv preprint arXiv …, 2018 - arxiv.org

… tion (eg, Wieting et al., 2016; Bojanowski et al., 2017) or learning separate **vectors** for each … form a context-independent token representation xk for each token position using pre-trained **word** em- beddings … This im- poses an inductive **bias** on the ELMo weights to stay close to an …

☆ Cited by 468 Related articles All 11 versions »

Word embeddings quantify 100 years of gender and ethnic stereotypes

N Garg, L Schiebinger, D Jurafsky… - Proceedings of the …, 2018 - National Acad Sciences

… **Word** embeddings quantify 100 years of gender and ethnic stereotypes. Nikhil Garg, Londa Schiebinger, Dan Jurafsky, and James Zou. PNAS April 17, 2018 115 (16) E3635-E3644; published ahead of print April 3, 2018 https://doi.org/10.1073/pnas.1720347115 …

☆ Cited by 45 Related articles All 10 versions

A context noise model of episodic **word** recognition.

S Dennis, MS Humphreys - Psychological review, 2001 - psycnet.apa.org

… that the participant would form a reinstated context that incorporates the context **vectors** from both … In single-**word** recognition, words of low normative frequency are recognized better than high-frequency … plotting hit rates against false alarm rates across a number of levels of **bias** …

☆ Cited by 379 Related articles All 19 versions

Word frequency effects in high-dimensional co-occurrence models: A new approach

C Shaoul, C Westbury - Behavior Research Methods, 2006 - Springer

… frequency information in a cal- culation, there is a danger that the highly skewed distribu- tion of **word** frequencies (a … will have much larger values in their vec- tors and many more nonzero elements in their **vectors**, they will … This “frequency **bias**” has been noted by Song et al …

☆ Cited by 58 Related articles All 9 versions

Convolutional neural networks for sentence classification

Y Kim - arXiv preprint arXiv:1408.5882, 2014 - arxiv.org

… by ci = f(w · xi:i+h-1 + b). (2) Here b ∈ R is a **bias** term and f is a non-linear function such as the hyperbolic tangent. This … labels. In one of the model variants, we experiment with having two 'channels' of **word vectors**—one that …

☆ Cited by 4028 Related articles All 18 versions »

Neural **word** embedding as implicit matrix factorization

O Levy, Y Goldberg - Advances in neural information processing …, 2014 - papers.nips.cc

… procedure is much more “symmetric”, in the sense that neither WW2V nor CW2V is orthonormal, and no particular **bias** is given to … The **word vectors** are evaluated by ranking the pairs according to their cosine similarities, and measuring the correlation (Spearman's ρ) with the …

☆ Cited by 864 Related articles All 13 versions »

Diachronic **word** embeddings reveal statistical laws of semantic change

WL Hamilton, J Leskovec, D Jurafsky - arXiv preprint arXiv:1605.09096, 2016 - arxiv.org

… The **word vectors** correspond to the rows of the matrix MPPMI ∈ R|V|×|VC | with entries given by MPPMI i,j = max { log (p(wi,cj) p(w)p(cj)) − α,0 } , (1) where cj ∈ VC is a context **word** and α > 0 is a negative prior, which provides a smooth- ing **bias** (Levy et al., 2015) …

☆ Cited by 203 Related articles All 18 versions »

Effects of high-order co-occurrences on **word** semantic similarity

B Lemaire, G Denhiere - Current psychology letters …, 2006 - journals.openedition.org

… In addition, it is probably not a strong **bias**: if co-occurrence relations in corpora reflect semantic information, they should appear in a … These association values were compared with the LSA cosine between **word vectors**: we selected the three best-ranked words as well as the …

☆ Cited by 89 Related articles All 18 versions »

Comparative experiments on disambiguating **word** senses: An illustration of the role of **bias** in machine learning

RJ Mooney - arXiv preprint cmp-lg/9612001, 1996 - arxiv.org

… A recent special issue of the Machine Learning journal on **Bias** Evaluation and Selec- tion … 2,859 binary features each representing the presence or absence of a particular **word** stem in … implementation of nearest neighbor was optimized to handle sparse bi- nary **vectors** by only …

☆ 99 Cited by 250 Related articles All 19 versions »»