Learning **word vectors** for sentiment analysis

AL Maas, RE Daly, PT Pham, D Huang, AY Ng… - Proceedings of the 49th …, 2011 - dl.acm.org

… vector representation φw = Rw corre- sponding to that **word's** column in R. The random variable θ is also a β-dimensional vector, θ ∈ Rβ which weights each of the β dimensions of words' representation **vectors**. We additionally introduce a **bias** bw for each **word** to capture …

☆ Cited by 1273 Related articles All 19 versions

Glove: Global **vectors** for **word** representation

J Pennington, R Socher, C Manning - Proceedings of the 2014 …, 2014 - aclweb.org

… Finally, adding an additional **bias** ˜bk for ˜wk restores the symmetry, wT i ˜wk + bi + ˜bk = log(Xik) … Because all unsupervised methods for learning **word vectors** are ultimately based on the occur- rence statistics of a corpus, there should be com- monalities between the models …

☆ Cited by 7213 Related articles All 28 versions

Man is to computer programmer as woman is to homemaker? debiasing **word** embeddings

T Bolukbasi, KW Chang, JY Zou… - Advances in neural …, 2016 - papers.nips.cc

… To quantify **bias**, we will compare a **word** vector to the **vectors** of a pair of gender-specific words. For instance, the fact that −−−! nurse is close to −−−−! woman is not in itself necessarily biased(it is also somewhat close to …

☆ Cited by 277 Related articles All 9 versions

A probabilistic model for semantic **word vectors**

AL Maas, AY Ng - NIPS Workshop on Deep Learning and …, 2010 - ai.stanford.edu

… vector representation φw = Rw corresponding to that **word's** column in R. The random variable θ is also a β-dimensional vector, θ ∈ R(β) which weights each of the β dimensions of words' representation **vectors**. We additionally introduce a **bias** bw for each **word** to capture …

☆ Cited by 52 Related articles All 7 versions

Related searches

word vectors **loss framework**

vector **space** word **representations**

word **representation global** vectors

word vectors **language modeling**

Tying **word vectors** and **word** classifiers: A loss framework for language modeling

H Inan, K Khosravi, R Socher - arXiv preprint arXiv:1611.01462, 2016 - arxiv.org

… W ∈ R|V |×dh and b ∈ R|V | are the the output projection matrix and the **bias**, respectively, and dh is the size of the RNN hidden … corresponds to the target **word** token (resulting in ut), and then take the inner product of the target **word** vector with all the other **word vectors** to get …

☆ Cited by 127 Related articles All 6 versions

Improving distributional similarity with lessons learned from **word** embeddings

O Levy, Y Goldberg, I Dagan - Transactions of the Association for …, 2015 - MIT Press

… (The last step follows because, as noted in Sec- tion 2, the **word** and context **vectors** are normal … The first- order terms measure the tendency of one **word** to appear in the context of the other … to PMI(w, c), while in GloVe it converges into their log-count (with some **bias** terms) …

☆ Cited by 684 Related articles All 12 versions

Document embedding with paragraph **vectors**

AM Dai, C Olah, QV Le - arXiv preprint arXiv:1507.07998, 2015 - arxiv.org

… Cluster analysis Generative model Dimensionality reduction Computational learning theory Functional decomposition Inductive **bias** Bayesian network … operations: pv("Lady Gaga") - wv("American") + wv("Japanese") where pv is paragraph **vectors** and wv is **word vectors** …

☆ Cited by 174 Related articles All 9 versions

Training restricted boltzmann machines on **word** observations

GE Dahl, RP Adams, H Larochelle - arXiv preprint arXiv:1202.5695, 2012 - arxiv.org

… ducing **word** representations and our learned n-gram features yield even larger performance gains … E(v, h) = −bv − ch − h Wv. (1) This energy is parameterized by **bias vectors** b ∈ RV and c ∈ RH and weight matrix W ∈ RH×V , and is converted into a probability distribution via …

☆ Cited by 90 Related articles All 17 versions

Compositional morphology for **word** representations and language modelling

J Botha, P Blunsom - International Conference on Machine Learning, 2014 - jmlr.org

… continuous space language models (CSLMs), an umbrella term for the LMs that represent words with real-valued **vectors**. Such **word** representations have been found to capture some morphological regular- ity (Mikolov et al., 2013b), but we … the language models' inductive **bias** …

☆ 99 Cited by 164 Related articles All 15 versions »

Skip-thought **vectors**

R Kiros, Y Zhu, RR Salakhutdinov, R Zemel… - Advances in neural …, 2015 - papers.nips.cc

… The computation is similar to that of the encoder except we introduce matrices Cz, Cr and C that are used to **bias** the update gate … The publically available CBOW **word vectors** are used for this purpose 2. The skip-thought models are trained with a vocabulary size of 20,000 words …

☆ 99 Cited by 979 Related articles All 15 versions »