*context2vec:*

**Learning Generic Context Embedding with Bidirectional LSTM**

**Oren Melamud**

Computer Science Dept. Bar-Ilan University melamuo@cs.biu.ac.il

**Jacob Goldberger**

Faculty of Engineering Bar-Ilan University goldbej@eng.biu.ac.il

**Ido Dagan**

Computer Science Dept. Bar-Ilan University dagan@cs.biu.ac.il

**Abstract**

Context representations are central to vari- ous NLP tasks, such as word sense disam- biguation, named entity recognition, co- reference resolution, and many more. In this work we present a neural model for efficiently learning a generic context em- bedding function from large corpora, us- ing bidirectional LSTM. With a very sim- ple application of our context represen- tations, we manage to surpass or nearly reach state-of-the-art results on sentence completion, lexical substitution and word sense disambiguation tasks, while sub- stantially outperforming the popular con- text representation of averaged word em- beddings. We release our code and pre- trained models, suggesting they could be useful in a wide variety of NLP tasks.

**摘要**

上下文表示是各种NLP任务的核心任务，例如语义解析，命名实体识别，同步参考解析等等。在本文工作中，我们提出了一个神经模型，使用双向LSTM从大型语料库中有效地学习通用语境的嵌入函数。我们设法在句子上通过我们的上下文表示来完成任务，且在词汇替换和词义消歧任务方面超越或接近达到最先进的结果，同时基本上超越了词汇的文本表示的最先进水平。我们的代码和预先训练的模型以及发布，被证明它们可以用于各种NLP任务。