**WORD REPRESENTATIONS VIA GAUSSIAN EMBEDDING**

**Luke Vilnis, Andrew McCallum**

School of Computer Science  University of Massachusetts Amherst  Amherst, MA 01003  luke@cs.umass.edu, mccallum@cs.umass.edu

**Abstract**

Current work in lexical distributed representations maps each word to a point vector in low-dimensional space. Mapping instead to a density provides many interesting advantages, including better capturing uncertainty about a representa- tion and its relationships, expressing asymmetries more naturally than dot product or cosine similarity, and enabling more expressive parameterization of decision boundaries. This paper advocates for density-based distributed embeddings and presents a method for learning representations in the space of Gaussian distributions. We compare performance on various word embedding benchmarks, investigate the ability of these embeddings to model entailment and other asymmetric relationships, and explore novel properties of the representation.

**摘要**

词法分布式表示中的当前工作是将每个单词映射到低维空间中的点向量。非密度的映射度提供了许多有趣的优势，包括更好地捕获关于表示及其关系的不确定性，比点积或余弦相似性更自然地表达不对称性，以及实现决策边界的更具表现力的参数化。本文提出了一种基于密度的分布式嵌入，并提出了一种在高斯分布空间中学习表示的方法。我们比较了各种词嵌入基准的性能，研究这些嵌入对蕴涵和其他不对称关系进行建模的能力，并探索表示的新特性。