

# We define a word as a vector

- Called an "embedding" because it's embedded into a space
- The standard way to represent meaning in NLP
- Fine-grained model of meaning for similarity
  - NLP tasks like sentiment analysis
    - With words, requires **same** word to be in training and test
    - With embeddings: ok if **similar** words occurred!!!
  - Question answering, conversational agents, etc

# Two kinds of embeddings

- **Sparse** (e.g. TF-IDF, PPMI)
  - A common baseline model
  - Sparse vectors
  - Words are represented by a simple function of the counts of nearby words
- **Dense** (e.g. word2vec)
  - Dense vectors
  - Representation is created by training a classifier to distinguish nearby and far-away words