# GloVe

## Technical Contributions

* A log-bilinear model that merges local context window methods with global matrix factorization
* Weighted least squares model trained on global word-word co-occurrence
* The goal of training is to learn word vectors such that their dot product equals the logarithm of the words' probability of co-occurrence.
* Training:
  + Build a co-occurrence matrix (X) with a fixed window size, where Xij represents how often the word i appears in the context of the word j
  + Uses co-occurrence ratios between two words in a context (assumes this is strongly connected to the meaning of the words)

## Strengths

* The model generates accurate word representations (75% performance on analogy task)
* Outperformed other methods on the word analogy task (Mikolov et al. (2013a))
* Uses global count statistics in addition to local word information

## Weaknesses

* Requires a lot of memory - needs dimension reduction
* Empirical results show that the performance of GloVe is very similar to that of word2vec, despite the different approaches of the models and the results showed in the paper (which show GloVe outperforming word2vec in both training times and accuracy)

## Improvements

* RGloVe: Uses cosine similarity between entity vectors instead of dot product to measure the entity occurrences, which more easily reaches a local optimum. Results on a corpus of Sina News show RGloVe outperforming GloVe.