Characterizing the impact of geometric properties of word embeddings on task performance

Brendan Whitaker, Denis Newman-Griffis, Aparajita Haldar
Hakan Ferhatosmanoglu, Eric Fosler-Lussier

Ohio State University
University of Warwick

June 4, 2019

Objective

Question

What geometric properties of an embedding space are important for performance on a given task?

Objective

Question

What geometric properties of an embedding space are important for performance on a given task?

- Understand utility of embeddings as input features.
- Provide direction for future work in training and tuning embeddings.

Embedding space?

In NLP, the term **embedding** is often used to denote both a map and (an element of) its image.

Definition

We define an **embedding space** as a set of word vectors in $\mathbb{R}^d$.

Geometric properties?

We consider the following attributes of word embedding geometry:

- position relative to the origin;
- distribution of feature values in $\mathbb{R}^d$;
- global pairwise distances;
- local pairwise distances.

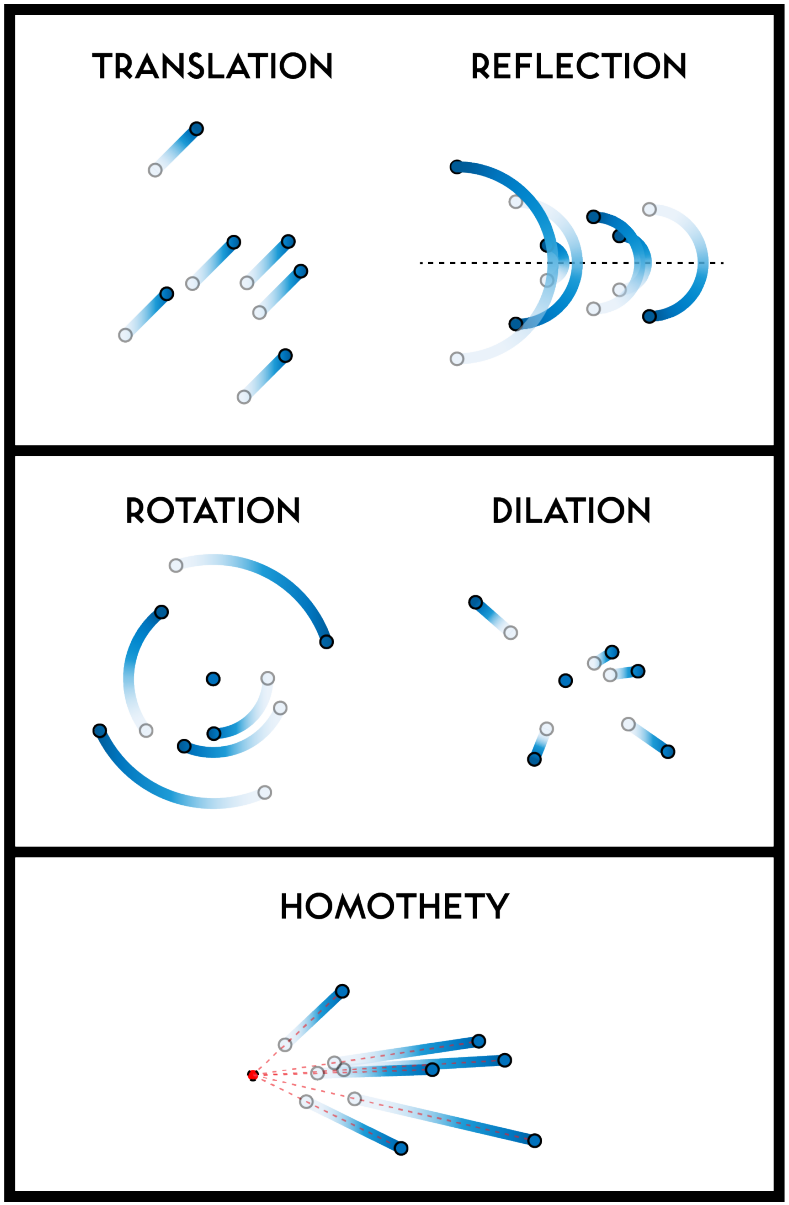
Our approach

Ablation Study

We transform the embedding space such that we expose only a subset of the stated properties to downstream models.

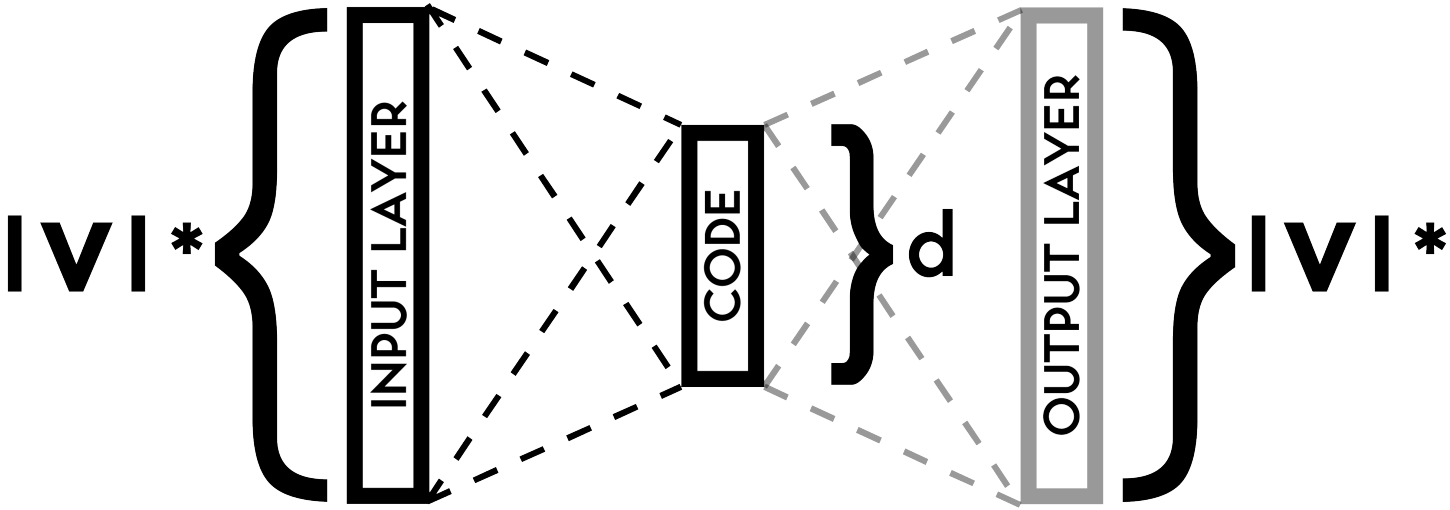
- position relative to the origin;
- distribution of feature values in $\mathbb{R}^d$;
- global pairwise distances;
- local pairwise distances.

Affine



- ~~pos. relative to the origin~~
- distribution of features
- global distances
- local distances

Cosine distance embedding (CDE)

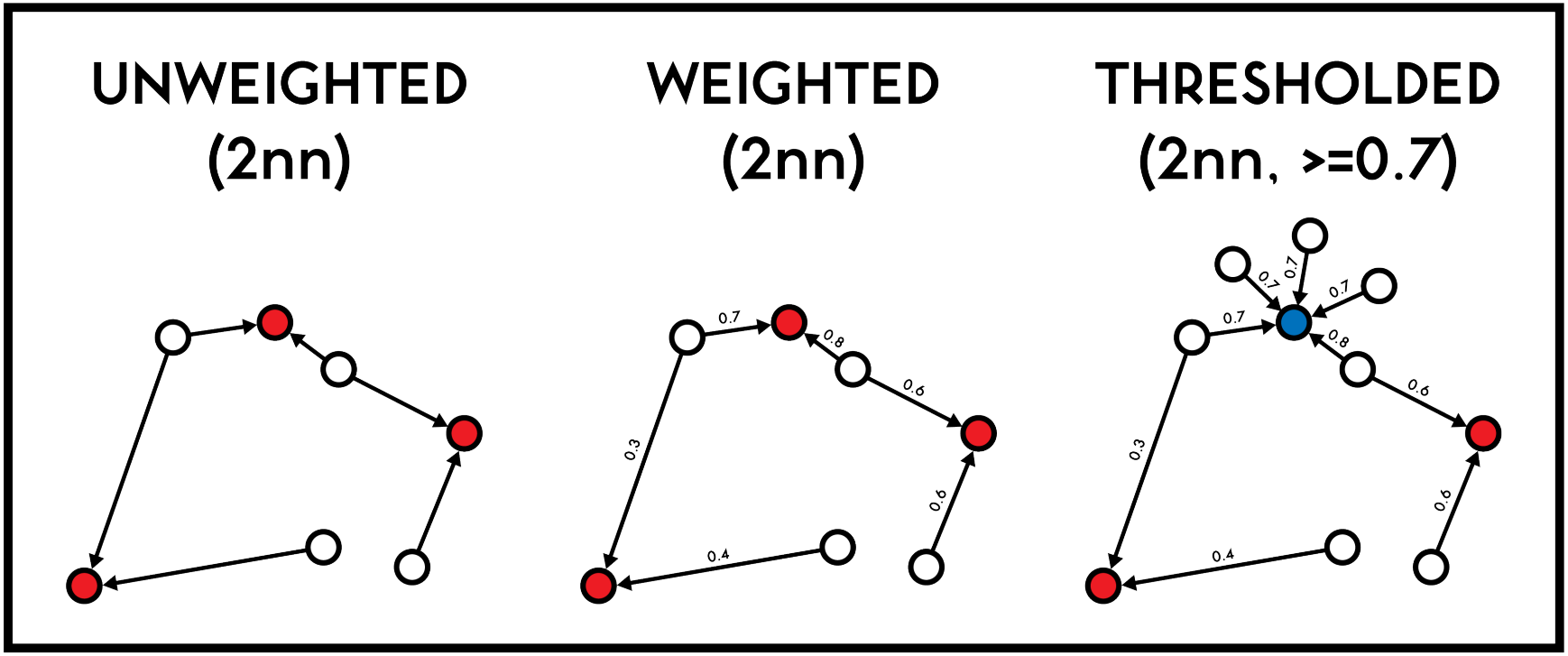


Specs:

- Activation function: ReLU;
- Epochs: 50;
- $d =$ embedding dimension (300);
- $|V|^* =$ distance vector dimension (10^4 most frequent words).

- ~~pos. relative to the origin~~
- ~~distribution of features~~
- global distances
- local distances

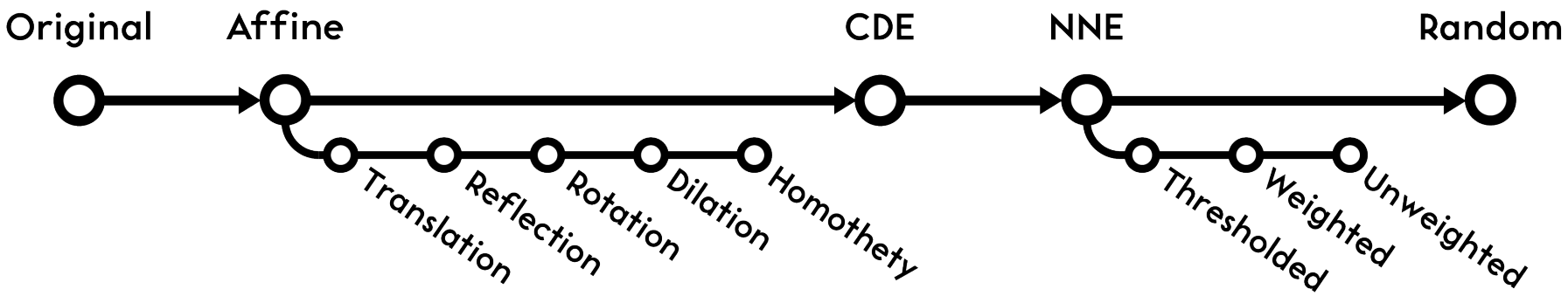
Nearest neighbor embedding (NNE)



Note: NN edges only drawn for colored nodes.

- ~~pos. relative to the origin~~
- ~~distribution of features~~
- ~~global distances~~
- local distances

Hierarchy of transformations



- Ordering is with respect to number of properties ablated.
- We include a random baseline of meaningless vectors.
- Arrow length does not mean anything.
- Transformations are applied independently to the original embeddings.

Embeddings and Tasks

Standard benchmark embeddings:

- Word2Vec on Google news;
- GloVe on common crawl;
- FastText on WikiNews.

Testing:

- 10 standard intrinsic tasks.
- 5 extrinsic tasks (embeddings plugged into a downstream machine learning model).

Tasks

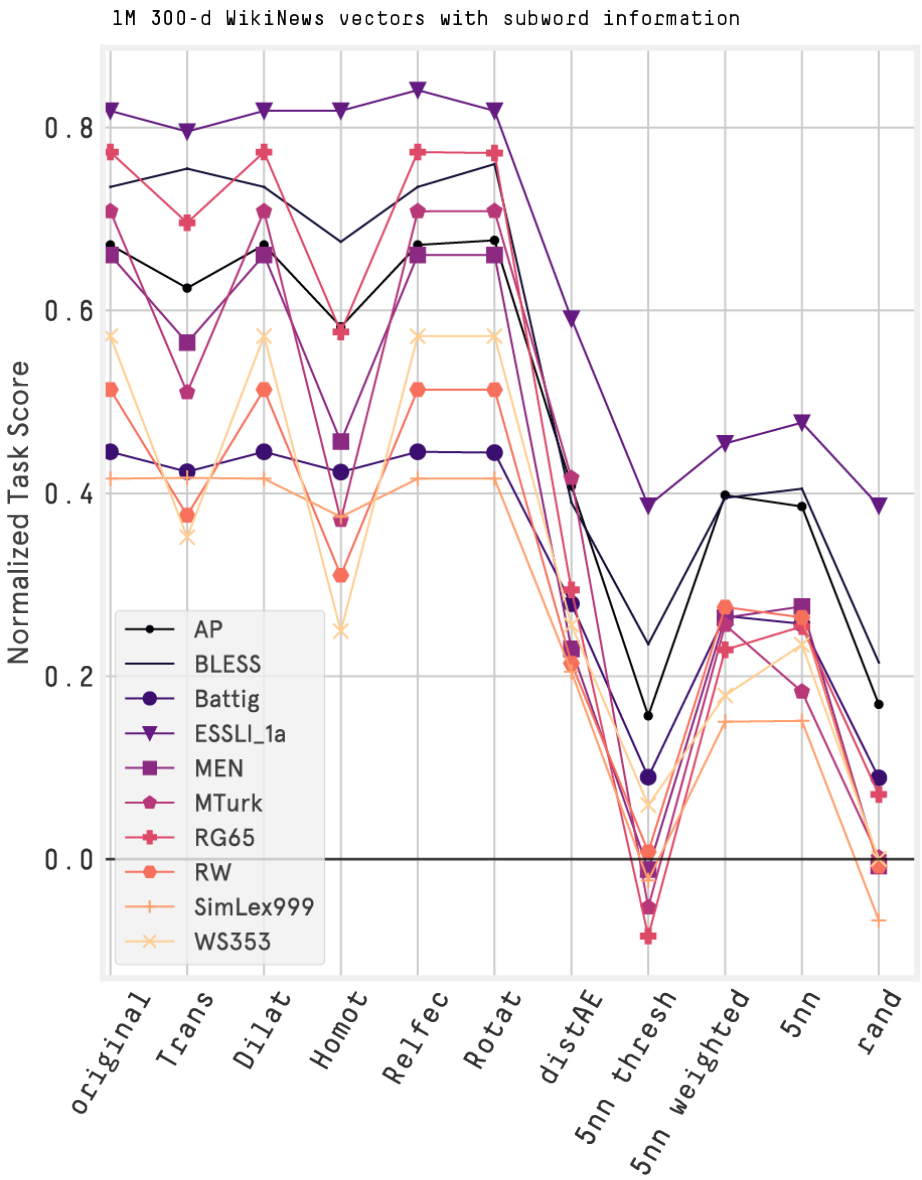
Intrinsic Tasks

- Word Similarity and Relatedness via cosine distance
  - WordSim353
  - SimLex-999
  - RareWords
  - RG65
  - MEN
  - MTURK
- Word Categorization
  - AP
  - BLESS
  - Battig
  - ESSLLI

Extrinsic Tasks

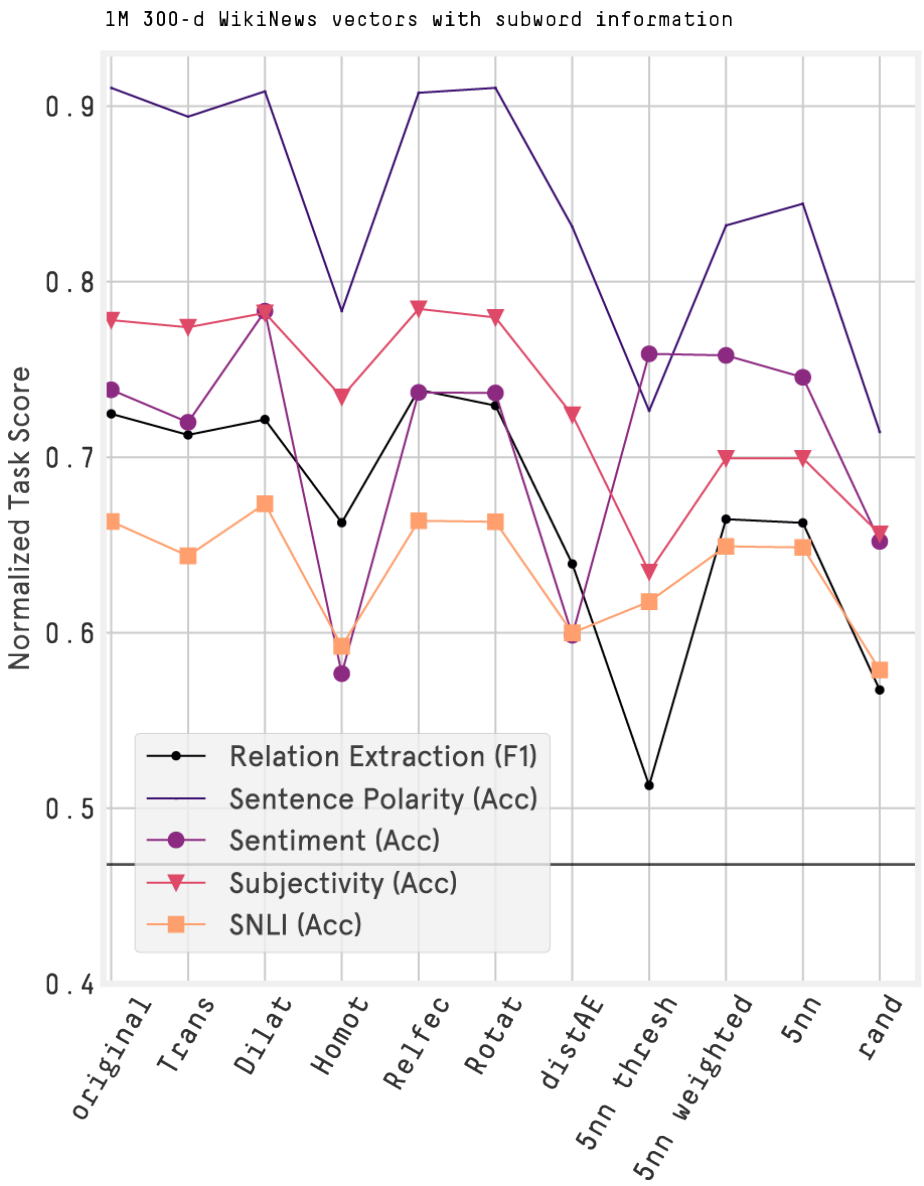
- Relation classif. on SemEval-2010 Task 8
- Sentence-level sentiment polarity classif. on MR movie reviews
- Sentiment classif. on IMDB reviews
- Subj./Obj. classif. on Rotten Tomatoes snippets
- SNLI

Results - intrinsic tasks



- We see the lowest performance on thresholded-NNE.
- Largest drop in performance at CDE (written as `distAE` on the graph).
- Rotations, dilations, and reflections are innocuous.
- Displacing the origin has a nontrivial effect.
- NNE causes a significant drop in performance as well.

Results - extrinsic tasks



- CDE is still the largest drop.
- NNE recover most of the losses, and are on par with affines.
- Extrinsic tasks are more robust to translations, but not homotheties.

Discussion

- Drop due to CDE likely associated with the importance of locality in embedding learning.
- With thresholded-NNE, high out-degree words are rare words, introducing noise during node2vec's random walk.

Takeaways

- We find that in general, both intrinsic and extrinsic models rely heavily on local similarity, as opposed to global distance information.
- We also find that intrinsic models are more sensitive to absolute position than extrinsic ones.
- Methods for tuning and training should focus on local geometric structure in $\mathbb{R}^d$.

Questions.

Questions?

github.com/OSU-slatelab/geometric-embedding-properties