Efficient Sentence Embedding via Semantic Subspace Analysis
Bin Wang, Fenxiao Chen, Yun-Cheng Wang and C.-C. Jay Kuo
University of Southern California (USC)

1. Highlight

Motivations:
- Efficiency: encoder large amounts of sentences
- Semantic grouping property of word embedding

2. Methodology

Step 1: Semantic Group Construction
- Leveraging Semantic grouping property of word vectors
- Handling variant-length input
- Weighted k-means algorithm

Step 2: Intra-group Descriptor
- Assign words into semantic groups
- Compute group center
- Calculate inter-group descriptor

Step 3: Inter-group Descriptor
- Model interaction between groups as sentence representation
- Vectorization

3. Application and Analysis

- Sentence Similarity
- Classification Tasks

4. Conclusions and Future Work

Future Work:
- With modularized design of S3E, we can try stronger clustering and correlation descriptors including subspace clustering, non-linear correlation computation with different kernel functions

Bibliography:

Code: https://github.com/BinWang28/SentenceEmbedding-S3E