Term Set Expansion based on Multi-Context Term Embeddings: an End-to-end Workflow

JiayiXu
2022/5/30

Term Set Expansion based on Multi-Context Term Embeddings: an End-to-end Workflow. Mamou et al. 2018

Content

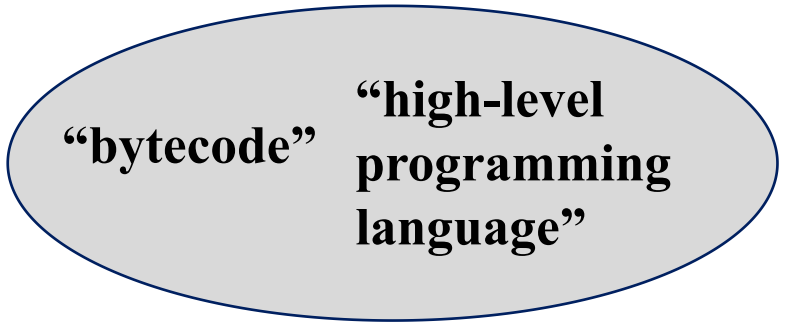
- **Section 1** Introduction
- **Section 2** Term Set Expansion Algorithm Overview
- **Section 3** System Workflow and Application
- **Section 4** Field Use Cases

What is Term set expansion?

topical similarity ⟺ **fuctional similarity**

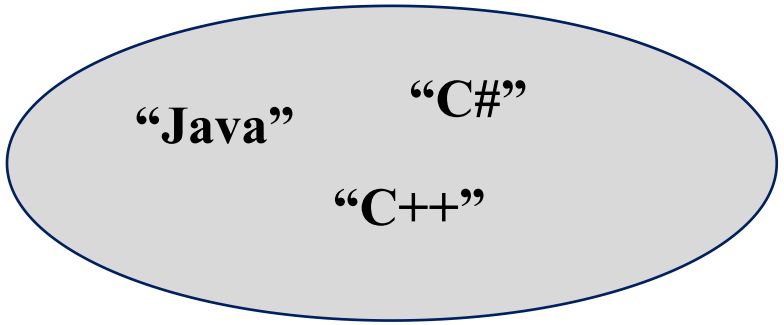
Python

linear bag-of-words contexts



Python

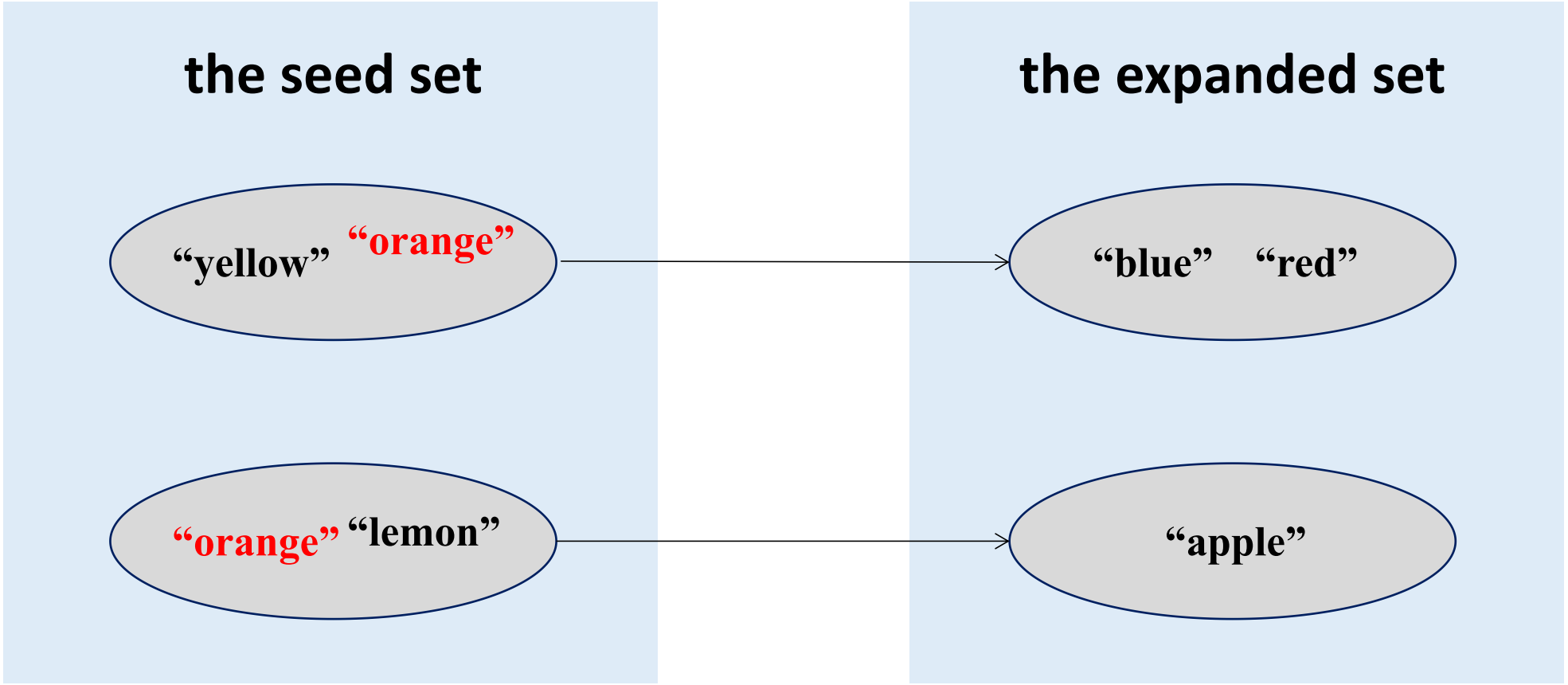
Term set expansion



Term Set Expansion based on Multi-Context Term Embeddings: an End-to-end Workflow. Mamou et al. 2018

Section 1 Introduction

Except for expansion from one single word, the terms of the expanded set are expected to be as functionally similar to all the terms of the seed set as possible.



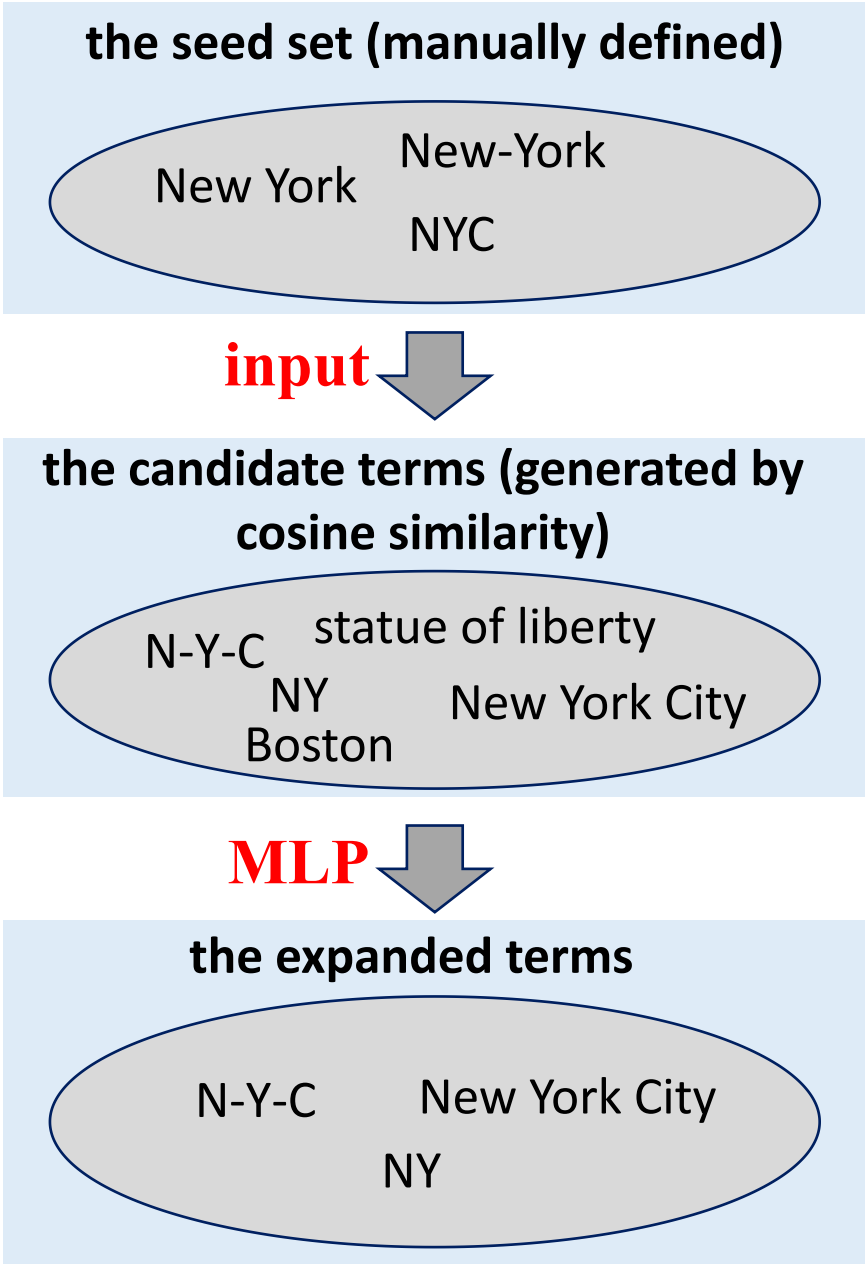
Term Set Expansion based on Multi-Context Term Embeddings: an End-to-end Workflow. Mamou et al. 2018

Section 2 Term Set Expansion Algorithm

Implementation

"For example, New York, New-York, NY, New York City and NYC are grouped together forming a single term group. Then, we use term groups as input units for embedding training; it enables obtaining more contextual information compared to using individual terms, thus enhancing the robustness of the embedding model."

"Taking the similarity scores between the seed terms and the candidate terms according to each of the different contexts as features, a Multilayer Perceptron (MLP) binary classifier predicts whether a candidate term should be part of the expanded set, where training and development term lists are used for the MLP training."



Term Set Expansion based on Multi-Context Term Embeddings: an End-to-end Workflow. Mamou et al. 2018