

基于分布式语义学的中文语义关系量化研究

林植楠

华中科技大学外国语学院, 湖北 武汉

收稿日期: 2024年7月23日; 录用日期: 2024年9月5日; 发布日期: 2024年9月20日

摘要

基于分布式语义学理论的词向量蕴含了丰富的语义信息, 一定程度上标志着自然语言处理和计算语言学领域进入了大模型发展时代。由于词向量的可计算属性, 逐渐发展出了多种基于词向量的语义计算任务, 语义关系辨析便是语义计算任务当中重要的一项。本研究基于fastText中文词向量和腾讯中文词向量的方法计算出表征语义关联强度的余弦相似度值, 并得出以下结论: fastText中文词向量和腾讯中文词向量在辨别近义关系、反义关系、上下义关系、部分-整体关系这4种语义关系的任务上表现存在一定差异; 通过比较Spearman相关系数, fastText中文词向量在实验数据上表现出其习得了更强的语义相似度特征, 腾讯中文词向量则体现出其学习到了更强的语义相关度特征; 在反义词辨析任务上, fastText中文词向量和腾讯中文词向量都在高度规约化的反义词对上计算出很高的余弦相似度值。

关键词

分布式语义学, 词向量, 语义关系计算, 语义相似性, 语义相关性

Quantitative Research on Chinese Semantic Relations Based on the Distributional Semantics

Zhinan Lin

School of Foreign Language, Huazhong University of Science and Technology, Wuhan Hubei

Received: Jul. 23rd, 2024; accepted: Sep. 5th, 2024; published: Sep. 20th, 2024

Abstract

The word embeddings, based on the distributed semantics theory, which contains rich linguistic information, have contributed a lot to the development of large language model (LLM) in the fields

of natural language processing and computational linguistics. Due to the computable properties of word embeddings, various semantic computing tasks based on them have gradually emerged, among which semantic relation discrimination is an important task in semantic computation. In our study, we adopt two word-embedding methods, the fastText Chinese word embeddings and the Tencent Chinese word embeddings, to calculate Chinese semantic relations, where the cosine similarity is used to represent the semantic association strength between words. The following are our findings in this study: First, the fastText Chinese embeddings and the Tencent Chinese embeddings show some differences in the task of distinguishing the four types of semantic relation in Chinese, namely, synonymy, antonymy, hyponymy and meronymy; Second, by comparing the Spearman correlation coefficient, the fastText embeddings have acquired more knowledge of semantic similarity between words, while the Tencent Chinese word embeddings have acquired more knowledge of semantic relatedness between words; Third, both the fastText Chinese embeddings and the Tencent Chinese word embeddings give higher values of cosine similarity to highly conventionalized antonyms.

Keywords

Distributional Semantics, Word Embeddings, Calculation of Semantic Relations, Semantic Similarity, Semantic Relatedness

Copyright © 2024 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

分布式语义学(Distributional semantics)起源于语言学中的结构主义理论[1],建立在 Harris (1954)提出的分布假设(Distributional Hypothesis)基础之上,强调语言的分布式属性[2]。分布式语义学认为语义上的相似性(similarity in meaning)会引起语言学分布上(即上下文语境等语言学特征)的相似性(similarity of linguistic distribution) [2]-[4],即在语料充足的情况下,语义相似的词会出现在相似的上下文语境中。Boleda 和 Erk 认为,相比结构主义语言学,分布式语义学舍弃了传统的用于表示语义的内在意蕴¹(intrinsic meaning)特征,将重心转移至了语义之间的几何数量关系上[5]。

随着自然语言处理技术的发展,分布式语义理论被应用在语义计算任务中,用向量的形式来表征语义,其大致流程为:将大量训练文本输入至计算机语言模型中,模型生成的词向量中蕴含了丰富的语义信息,也被称为语义向量空间(semantic vector space),其中意义相近的词对应的词向量会在语义向量空间中相距很近,意义不同的词则会在向量空间里相隔更远[6]。Boleda 认为自然语言处理领域之所以采用分布式语义学来对语义建模,有以下三点原因:1) 语义的分布式特征从自然语言数据中习得,因此极具实证性[6]。此外,分布式语义建模可以让模型从大量语料中自动学习语义信息,在语料充足的情况下还可以跨语言进行训练,有助于研究人员系统性研究不同语言的语义特征。2) 基于分布式语义学获取的词向量具有高维性的特点,即从自然语言数据中提取的丰富且细微的语义特征蕴藏在词向量的数百维度中,而传统的结构主义方法的可操作性十分受限,只可提取极少数语义特征,难以表征出不同词语的大量语义特征。3) 基于分布式语义理论的词向量由向量组成,因此可以用词向量计算的方法对语义进行量化研究。

正如上文所说,由于词向量具备可运算性,于是自然语言处理领域和计算语言学领域有许多语义计

¹内在意蕴(intrinsic meaning): 指一个词语或短语本身所固有的含义,与外部因素无关。

算任务和方法应运而生，语义相似度计算便是其中最为重要的一种。然而，用词向量相似度方法计算出来的结果究竟反映了语义相似性(semantic relatedness)还是语义相关性(semantic relatedness) [7]？就拿反义词而言，反义词之间的语义相似度低，然而反义词之间的语义相关度高，那使用词向量方法计算反义词之间的语义相似度时会有何种表现呢？本文将通过使用词向量计算的方式来计算属于不同语义联系类型的词语之间的余弦相似度值，用以探究词向量计算结果表征的是语义之间的相似性还是相关性，以及通过实验研究词向量计算出来的结果是否与人类的判断结果具有相关性，最后通过实验来评估词向量方法在反义词辨析任务上的表现。

2. 实验设计

2.1. 研究对象

本文将通过基于词向量的方法对汉语中不同的语义关系进行研究，将分别研究中文语义关系的三个方面：1) 词向量计算方法对属于不同语义联系类别的词语间的区分度；2) 在人为给定词语的相关性/相似性得分的数据集基础之上，比较词向量计算值与人工标注值之间的相关性；3) 词向量计算对于中文反义词的表征情况。

2.2. 数据集

2.2.1. 语义联系类型分类数据集

本研究使用了 NLPPCC 2017 Shared Task 中用于研究中文语义关系类型(Chinese Word Semantic Relation Classification)的数据集，用词向量计算方法来可视化词向量对不同语义联系类别的表征情况。主办方人工标注了隶属于四种不同语义联系类别的中文数据，每种语义联系类别各包含 500 对词对，包括名词、动词、形容词，以及副词、连词等功能词。数据集中包含的四种词汇语义联系类别分别是近义关系(Synonym, 人工标记为 s)、反义关系(Antonym, 人工标记为 a)、上下义关系(Hyponym, 人工标记为 h)、部分 - 整体关系(Meronym, 人工标记为 m) [8]：

- 1) 近义关系(Synonym)：词语 A 与词语 B 的意义相似，e.g., 色彩缤纷 vs. 姹紫嫣红；
- 2) 反义关系(Antonym)：词语 A 与词语 B 的意义相反，e.g., 骄傲 vs. 谦虚；
- 3) 上下义关系(Hyponym)：词语 B 是词语 A 所指范畴下的一种，e.g., 菊花 vs. 花；
- 4) 部分 - 整体关系(Meronym)：词语 B 是词语 A 的一部分，e.g., 花蕾 vs. 花。

本实验剔除掉 fastText 和腾讯词向量词表中所不包含的字词后，最终剩余 447 对近义词对，483 对反义词对，445 对上下义关系词对，429 对部分 - 整体关系词对。

2.2.2. 语义相似度/相关度人工打分数据集

为探究词向量表征语义关系与人类认知语义关系两者之间的异同，本实验采用两个人工标注的带有语义相似度/关联度分数的数据集用于研究：

1) Wordsim_501 数据集：第一份数据由 wordsim-240 和 wordsim-297 两份数据集组合而成，这两个数据集分别包含 240 对中文词语和 297 对中文词语，每对词语都被人工标注了词语间的语义相关性分数(relatedness score)，分数越高，表示人们认为两个词语之间的语义关联度越高，如“李白”和“诗”两个词之间的相关度分数很高，而“蛋白质”和“文物”两个词之间的相关度分数很低[9]。由于部分词语不包含在 fastText 词向量和腾讯词向量的词表中，所以去除该部分词语后，可实际用于本研究的词对共 501 条。

2) Miller 数据集：来自 Miller 等人的研究，是为研究语义相似度与语境相似度之间的关系而设计

[10]。这份数据中包含 30 对英文词对, 在国内研究中, 杨泉在保留语义相似度值的前提下将这三十对英文词翻译为了中文, 本文将采用中文译版对其进行研究[11]。每个词对的两个词之间的语义相似度值(similarity score)由人工标注, 分数越高, 代表两个词语的语义越相似、词对内的两个词语可相互替换的程度越高, 如“轿车”和“汽车”语义十分相似, 所以得分高, 而“中午”和“绳子”的语义差异很大, 因此得分低。在剔除了不包含在 fastText 和腾讯词向量词表中的词语后, 共有 29 条词对用于本研究。

2.2.3. 反义词数据集

为研究词向量对中文反义关系的语义理解情况, 本研究从《新华反义词典》中随机抽取了部分反义词对, 如: 年初 - 年底[12]; 同时, 实验还采用了已有研究中设计的存在反义关系的单个汉字间组成的反义字对, 如: 南 - 北, 男 - 女[7]。本文将通过词向量计算的方法对所有的字词对进行余弦相似度计算, 从而获取反义词对之间的语义关联/相似分数。

2.3. 实验方法

本文使用预训练好的 fastText 中文词向量[13]以及腾讯中文词向量[14]计算出每个中文词汇的词向量, 再通过计算出词向量之间的余弦相似度值(cosine similarity)表示词语间的语义相似度/相关度, 从而探究 fastText 词向量和腾讯中文词向量对词汇间语义关系的表征情况, 本文共进行三个实验:

对于语义联系类型分类数据集, 我们将计算出来的词对间的余弦相似度值可视化, 用直观的方式呈现词向量对不同语义联系类别的表征有何差异。

对于语义相似度/相关度人工打分数据集, 我们将会对实验得到的余弦相似度值与人工打分分值进行 Spearman 相关系数计算, 用以评估词向量对词汇间语义关系的抓捕与人类对语义关系理解之间的异同。一般来说, Spearman 相关系数数值越高, 代表两个数据之间越存在显著的相关性, 即代表词向量方法就越可能吻合人类对于词语间语义关系的判断模式。

对于反义词数据集, 本实验将基于余弦相似度值, 通过结果分析的方式来评估词向量的方法能在何种程度上表征出反义关系。

3. 实验结果与分析

3.1. 实验一: 不同语义联系类型的余弦相似度值计算

在获得词向量之后, 通过计算对应不同词的词向量之间的余弦相似度值, 并对余弦相似度值可视化, 从而得到如图 1(a)、图 1(b)和图 2 中的结果。

由图 1(a)和图 1(b)可以得知, 对于反义关系、近义关系、部分 - 整体关系、上下义关系这四种不同类别的语义关系, 腾讯词向量计算出来的词语之间的余弦相似度值总体而言比 fastText 词向量计算出来的值高, 可见两种词向量由于训练模型的原理不同、参数不同、训练语料不同等原因, 会对属于同一语义类型的同一数据得出不同的计算结果。

由图 2 可以看出, 在使用两种不同的词向量进行余弦相似度计算的情况下, 对应不同语义关系类别的词向量计算出来的余弦相似度值有着不同的分布。

就总体密度分布而言, 基于 fastText 中文词向量和腾讯中文词向量计算出来的余弦相似度值都未能在不同语义关系类别间呈现显著差异, 尤其是在近义关系和反义关系的分辨上并未拉开明显区分度。可见基于这两种词向量计算出来的余弦相似度值都不能很好地区分开各语义联系类别。

就余弦相似度值的中位数而言。

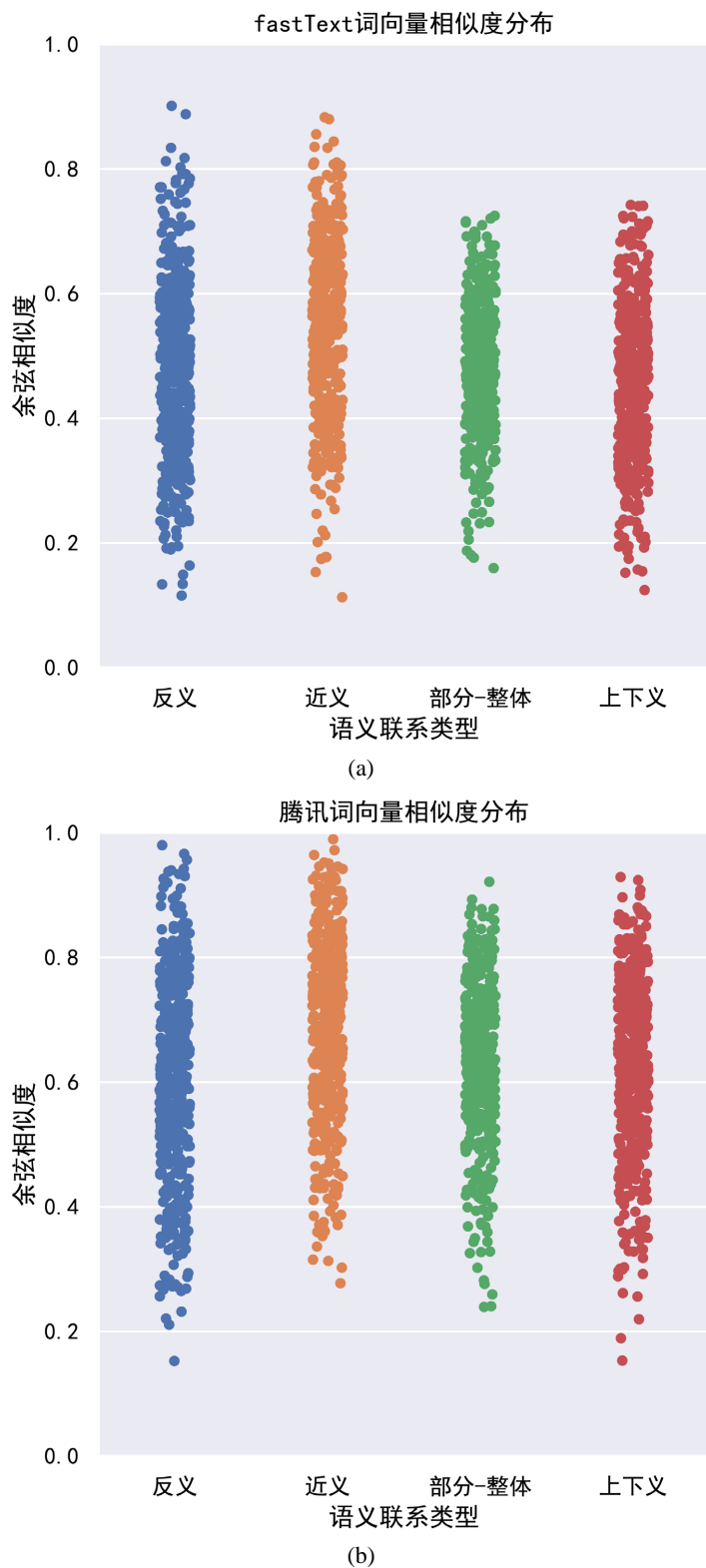


Figure 1. (a) Diagram of the fastText embeddings on discriminating four types of semantic relations; (b) Diagram of the Tencent embeddings on discriminating four types of semantic relations

图 1. (a) fastText 词向量计算四种语义联系类型的余弦相似度值图；(b) 腾讯词向量计算四种语义联系类型的余弦相似度值图

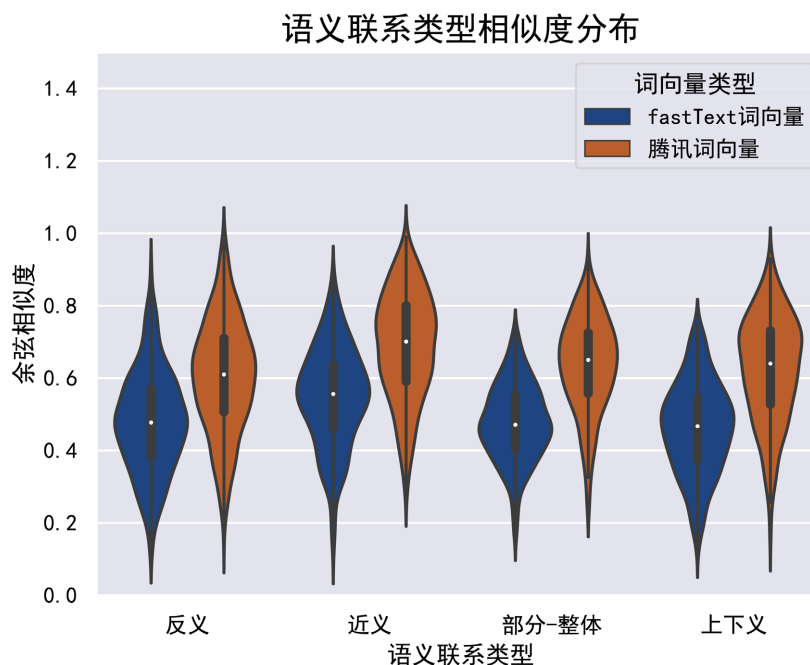


Figure 2. Diagram of the differences between the fastText embeddings and Tencent embeddings on discriminating four types of semantic relations

图 2. fastText 词向量和腾讯词向量余弦相似度值计算结果对比图

对于 fastText 中文词向量, 近义关系 > 反义关系 \approx 部分 - 整体关系 \approx 上下文关系;
对于腾讯中文词向量, 近义关系 > 部分 - 整体关系 > 上下文关系 > 反义关系。

3.2. 实验二：词向量计算与人工评分间的 Spearman 相关系数

在 Wordsim_501 数据集上, 基于 fastText 词向量获得的余弦相似度值与人工打分之间的 Spearman 相关系数为 0.572; 基于腾讯词向量获得的余弦相似度值与人工打分之间的 Spearman 相关系数为 0.582。

由表 1 可以得出, 在 Miller 数据集上, 基于 fastText 词向量获得的余弦相似度值与人工打分之间的 Spearman 相关系数为 0.862; 基于腾讯词向量获得的余弦相似度值与人工打分之间的 Spearman 相关系数为 0.578。

Table 1. The Spearman correlation coefficient between the cosine similarity and the human rating scores

表 1. 余弦相似度值与人工打分间的 Spearman 相关系数表

数据集	fastText 词向量	腾讯词向量
Wordsim_501 数据集	0.572	0.582
Miller 数据集	0.862	0.578

基于此实验结果, 我们发现 fastText 词向量计算出来的余弦相似度值在 Miller 数据集上获得了更高的 Spearman 系数(0.862)远高于 Wordsim_501 数据集的结果(0.572), 这意味着 Miller 数据集中的人工打分和通过 fastText 词向量计算出来的余弦相似度值之间存在更显著的相关关系, 也就是说, fastText 词向量的结果主要体现了词语语义之间的相似性。

与之相反的是, 腾讯词向量计算出来的余弦相似度值在 Wordsim_501 数据集上获得了更高的 Spearman 系数(0.582), 这意味着腾讯词向量的计算结果主要体现了词语语义之间的相关性。

3.3. 实验三：中文反义字/词余弦相似度计算

由于词向量在表征反义关系任务上一直表现欠缺，因此本文分别选取了余弦相似度计算值最高的 20 组反义词对和 20 组反义字对，通过具体的案例分析来探究为何基于分布式语义学的词向量方法会给予这些反义字/词极高的余弦相似度值，实验结果分别见表 2 和表 3。

Table 2. The cosine similarity of Chinese words antonyms

表 2. 反义词对词向量余弦值计算

序号	fastText 词向量 Top20 词对		余弦值	腾讯词向量 Top20 词对		余弦值
1	男方	女方	0.917822	南边	北边	0.99036
2	男人	女人	0.885754	男方	女方	0.989191
3	买方	卖房	0.878317	买方	卖方	0.982858
4	男性	女性	0.862576	男声	女声	0.96726
5	南边	北边	0.854464	南方	北方	0.964806
6	男声	女声	0.845091	左翼	右翼	0.962998
7	今年	去年	0.837644	男人	女人	0.948912
8	南方	北方	0.826936	荤菜	素菜	0.947684
9	明枪易躲	暗箭难防	0.82261	男性	女性	0.938402
10	弄瓦	弄璋	0.817134	肉中刺	眼中钉	0.934858
11	男子	女子	0.817064	男式	女式	0.930357
12	那儿	这儿	0.808386	真话	假话	0.926496
13	一言丧邦	一言兴邦	0.806793	治本	治标	0.922112
14	递增	递减	0.804579	以进为退	以退为进	0.916445
15	年初	年底	0.803302	看跌	看涨	0.913792
16	冬天	夏天	0.800843	实词	虚词	0.912499
17	左翼	右翼	0.79254	轻型	重型	0.909057
18	那个	这个	0.78957	冬天	夏天	0.907086
19	实症	虚症	0.787618	略胜一筹	稍逊一筹	0.901871
20	冷色	暖色	0.782409	进攻	防守	0.896219

表 2 的结果表明，从词义的维度上来看，反义词之间往往只在某个语义维度上存在显著差异，而在大部分语义维度上极为相似，而词向量往往维度很高，没法突出反义词词义间的显著差异项；从篇章维度来看，尽管反义词意义相反，但其出现的语境却极有可能相似，比如“以退为进”和“以进为退”，“那个”和“这个”等，错误地符合了分布式假设对于语义相似度的定义，这都对基于分布式语义理论的词向量辨析反义词带来了挑战。

此外，对于“明枪易躲，暗箭难防”、“治标不治本”这类由意义相反的词组合而成的固定表达，它们有极高的频率会同时出现在同一句语言表达中，这也就导致尽管“明枪易躲”和“暗箭难防”、“治标”和“治本”是意义相反的词，但它们出现的上下文语境基本一致，于是基于分布式语义理论的词向量会认为这两个表达在语义上具有很高的相似性。

Table 3. The cosine similarity of Chinese single characters antonyms
表 3. 反义字对词向量余弦值计算

序号	fastText 词向量 Top20 字对		余弦值	腾讯词向量 Top20 字对		余弦值
1	右	左	0.935866	女	男	0.946104
2	北	南	0.834038	北	南	0.942414
3	我	你	0.767933	胖	瘦	0.894073
4	女	男	0.767543	大	小	0.890323
5	买	卖	0.746019	东	西	0.877673
6	阴	阳	0.744586	粗	细	0.873199
7	嫁	娶	0.731184	横	竖	0.858366
8	秋	春	0.724572	阴	阳	0.850391
9	白	黑	0.708508	那	这	0.848983
10	东	西	0.698332	强	弱	0.848094
11	三	一	0.692905	旱	涝	0.841871
12	低	高	0.681713	软	硬	0.833566
13	初	末	0.674881	白	黑	0.829863
14	饱	饿	0.669912	低	高	0.824796
15	红	绿	0.663662	红	绿	0.823128
16	快	慢	0.658425	薄	厚	0.818868
17	输	赢	0.649404	近	远	0.817501
18	薄	厚	0.643431	买	卖	0.816125
19	冬	夏	0.640223	秋	春	0.814542
20	横	竖	0.639381	嫁	娶	0.812384

表 3 中记录了余弦相似度值最高的 20 组反义字对, 我们可以从这 20 组的数据中观察发现, 由于这些反义词表达是日常生活中高频使用的一些基本概念表达, 它们出现的上下文语境也往往极为相似, 如“左”和“右”往往表示方位(“图书馆在左/右边”)、“快”和“慢”表示速度(“他跑步的速度很快/慢”)。对于基于分布式语义学的词向量而言, 会认为上下文语义相似的词语义也会相似, 因此这会给词向量正确识别出反义词带来很大的挑战。

此外, 从历时角度来看, 这些字中有许多反义字间已组合成为了中文中的固定词组, 又称反义相成词。反义相成词指的是由两个在意义上相反或相对的单音节词素组合而成的合成词, 如“左右”、“男女”、“胖瘦”、“买卖”、“阴阳”、“东西”等词组[15]。反义相成词往往存在意义虚化的特征, 即整体词义超出了两个反义词素(如: 左和右)的意义之和, 发展出了更具普遍性的意义, 在现实生活中, “左右”一词并非表示方位, 而是表示更为抽象的“大约”的语义。反义相成词的形成可以从侧面说明这些反义字往往出现在高度相似的上下文语境当中, 甚至组合成日常中使用频率十分高的反义相成词, 因此对于基于分布式语义理论发展出来的词向量来说, 识别出这些高度规约化的反义词变得极具挑战性。

同时, Aina 等(2020)认为尽管我们可以通过词向量方法对语义进行几何计算, 但基于分布式语义学获得的词向量是对词语出现的所有语境情况进行抽象处理, 往往表征的是一个词语普遍上的语义, 并不

像人类思维一样灵活到足以对词语所处的每一种具体的语境作出反应[16][17],因此尽管词向量可以习得语言数据当中的大量的语义信息,但远比不上人类认知可以达到的程度。

4. 结论

本文可以得出以下三个结论:

1) 在计算汉语中的反义、近义、上下文和部分-整体关系这四种语义联系类型的任务上,fastText词向量和腾讯词向量的实验结果存在一定差异;

2) 基于词向量的方法判断两个词语间的语义相似度时,腾讯词向量的生成原理决定了其计算结果主要体现了两个词语间的相关性而非相似性,fastText词向量的计算结果更多地体现出了两个词语间的相似性而非相关性;

3) 此外,词向量更多的代表的是语义上的共性而非个性,如果反义词出现的上下文语境差不多,那这对反义词对应的词向量会很相似,这是导致词向量在反义词辨析任务上表现欠佳的原因之一。

因此,笔者认为,为提升语言模型在复杂多样的语义关系处理任务上的能力,离不开对自然语言处理下游任务的进一步微调,如通过推理训练、投喂不同语义类型的语料等使语言模型具备更加专精的能力,让语言模型在某些自然语言任务上的结果变得更加类人。最后,笔者认为,由于人类对语义的理解与表达离不开对物理世界的感知,因此引入多模态这类非语言特征是否会对语言模型表征语义关系产生积极影响也有待进一步研究。

致 谢

在此真挚地感谢我的导师唐旭日教授对此篇研究给予的悉心指导。

参考文献

- [1] Sahlgren, M. (2006) The Distributional Hypothesis. *The Italian Journal of Linguistics*, **20**, 1-18.
- [2] Harris, Z.S. (1954) Distributional Structure. *WORD*, **10**, 146-162. <https://doi.org/10.1080/00437956.1954.11659520>
- [3] Firth, J.R. (1957) A Synopsis of Linguistic Theory 1930-1955. *Studies in Linguistic Analysis*, **Special Volume**, 10-32.
- [4] Rubenstein, H. and Goodenough, J.B. (1965) Contextual Correlates of Synonymy. *Communications of the ACM*, **8**, 627-633. <https://doi.org/10.1145/365628.365657>
- [5] Boleda, G. and Erk, K. (2015) Distributional Semantic Features as Semantic Primitives—Or Not. *AAAI Spring Symposia*, **3**, 2-5.
- [6] Boleda, G. (2020) Distributional Semantics and Linguistic Theory. *Annual Review of Linguistics*, **6**, 213-234. <https://doi.org/10.1146/annurev-linguistics-011619-030303>
- [7] Zhang, Z., Yuan, P. and Jin, H. (2023) Exploring Word-Sememe Graph-Centric Chinese Antonym Detection. In: Koutra, D., Plant, C., Rodriguez, M.G., Baralis, E. and Bonchi, F., Eds., *Machine Learning and Knowledge Discovery in Databases: Research Track*, Springer, 583-600. https://doi.org/10.1007/978-3-031-43418-1_35
- [8] Wu, Y. and Zhang, M. (2018) Overview of the NLPCCC 2017 Shared Task: Chinese Word Semantic Relation Classification. In: Huang, X.J., Jiang, J., Zhao, D.Y., Feng, Y.S. and Hong, Y., Eds., *Natural Language Processing and Chinese Computing*, Springer, 919-925. https://doi.org/10.1007/978-3-319-73618-1_81
- [9] Chen, X., Xu, L., Liu, Z., Sun, M. and Luan, H. (2015) Joint Learning of Character and Word Embeddings. *Proceedings of the Twenty-Fourth International Joint Conference on Artificial Intelligence*, Buenos Aires, 25-31 July 2015, 1236-1242.
- [10] Miller, G.A. and Charles, W.G. (1991) Contextual Correlates of Semantic Similarity. *Language and Cognitive Processes*, **6**, 1-28. <https://doi.org/10.1080/01690969108406936>
- [11] 杨泉. 知识本体与词向量结合的词义相似度强化学习计算方法[J]. 重庆理工大学学报(自然科学), 2022, 36(1): 238-235.
- [12] 张志毅, 张庆云. 新华反义词词典(中型本)[M]. 上海: 商务印书馆, 2008.

- [13] Bojanowski, P., Grave, E., Joulin, A. and Mikolov, T. (2017) Enriching Word Vectors with Subword Information. *Transactions of the Association for Computational Linguistics*, **5**, 135-146. https://doi.org/10.1162/tacl_a_00051
- [14] Song, Y., Shi, S., Li, J. and Zhang, H. (2018) Directional Skip-Gram: Explicitly Distinguishing Left and Right Context for Word Embeddings. *Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 2 (Short Papers)*, New Orleans, 8 June 2018, 175-180. <https://doi.org/10.18653/v1/n18-2028>
- [15] 谷鸿洁. 反义相成词研究综述[J]. 现代语文, 2015(7): 10-12.
- [16] Aina, L., Gulordava, K. and Boleda, G. (2019). Putting Words in Context: LSTM Language Models and Lexical Ambiguity. *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*, Florence, 28 July-2 August 2019, 3342-3348. <https://doi.org/10.18653/v1/p19-1324>
- [17] McNally, L. and Boleda, G. (2017) Conceptual versus Referential Affordance in Concept Composition. In: Hampton, J.A. and Winter, Y., Eds., *Compositionality and Concepts in Linguistics and Psychology*, Springer, 245-267. https://doi.org/10.1007/978-3-319-45977-6_10