

南 京 大 学

本科生毕业论文（设计、作品）指导教师评阅意见

指导教师评语：（不少于 300 字）

词汇语义演变与社会文化的变迁息息相关，能够服务于跨领域理论以及实践研究。该论文以《人民日报》上 1946-2022 年的新闻数据作为语料库，并使用历时词嵌入模型，对语义演变进行定性和定量分析，该研究方向具有一定的理论与实用价值。

该论文使用基于 Word2Vec 的 TWEC 模型，在 1946-2022 年的《人民日报》新闻数据上对不同粒度的时间片训练中文历时词嵌入模型。获得词表征后，本文筛选出可能存在语义变化的词表，分别进行了定性和定量分析。此外，本文分别从角色变化和概念变化两个方面设计了中文的跨时间类比任务，对词向量进行测试，获得的结果与现实情况能够匹配。最后，尝试用跨时间类比词捕捉流行病的发展趋势，并提出历史词嵌入模型用于与时事有关的社会学研究的一种可能路径。

该论文结构合理，逻辑清晰，语句通顺，达到本科学位论文要求，同意参加本科论文答辩，并建议授予学士学位。

指导教师签名：



2023 年 5 月 22 日

南 京 大 学

本科生毕业论文（设计、作品）评阅教师评阅意见

评阅教师评语：（不少于 300 字）

该论文以 1946-2022 年《人民日报》的新闻数据作为语料库，使用历时词嵌入模型，对语义演变进行定性与定量分析，该研究课题能够服务于跨领域的理论研究与事件预测等，具有一定的理论与实用价值。

文章使用 TWEC 模型在语料上对不同粒度的时间片训练中文历时词嵌入模型。随后，对存在语义变化的词表进行定量与定性分析。通过单词邻域来判断语义变化，绘制了词向量的演变轨迹，并对语义变化的速度进行量化表示和分析。此外，文章分别从语言内部变化和语言对外部世界变化的反映两个角度作了讨论，利用角色和概念的跨时间类比任务证实历时词表征在捕捉现实事件方面的可行性。

该论文结构科学，逻辑清晰，内容完整，观点突出，达到本科学位论文要求，同意参加本科论文答辩，并建议授予学士学位。

评阅教师签名：高坤

2023 年 5 月 24 日

南 京 大 学

本科生毕业论文（设计、作品）答辩记录、成绩评定

答辩记录：

1. 问题：为什么如表 17 的标题是“微信”而不是“通信方式”的类比词？

答：这里以“微信”为参照，如 2022 年的微信相当于 1946 年的信件、电报。
（表名有歧义，“‘微信’（2022 年）”不易看出具体类比的对象，建议修改。）

2. 问题：除了每个词的含义发生什么变化，是否存在一些群体性的现象，比如
在某一时期所有词的含义都发生了较大变化，或者发生较大变化的词尤其多？

答：之前没有考虑过这个方面，可以再进行一些探索。

3. 问题：这篇论文最后的结论是什么？

答：这方面目前大多数工作都是在英文上进行的，我主要在中文上做了一些初步
的尝试，说明历时词嵌入模型也能发现中文词义演变和用于研究中国社会。

答辩记录人签名：徐思厚

答辩小组评语：（不少于 100 字）

该同学论文选题新颖、工作扎实，行文逻辑清晰、格式规范，答辩中思路清
楚、重点突出，能够准确回答与工作相关的问题，表现出在本学科的基础理论和
专业知识。答辩小组一致同意通过该同学的本科答辩，并建议授予学士学位。

答辩小组成员：

顾坤 黄超 任建

成绩

95

组长签名：

顾坤

答辩时间：2023 年 5 月 26 日