

周报

屈原斌

2020 年 10 月 19 日

1 上周任务

- 看论文;
- 论文综述 (未完成);

1.1 任务详细

- 对已有论文中的方法按照不同维度进行划分:
 - 按照模型学习方法维度进行划分 (监督学习/无监督学习), 结果见表1;
 - 按照文本表示方法维度进行划分 (概率分析/语义分析/机器学习), 结果见表2:
 - * 概率分析: 使用文本表面特征和概率估计, 例如CVA相似度计算;
 - * 语义分析: 基于语料库 (LDA, RI方法) 或基于词表 (WordNet) 的文本相似性计算方法;
 - * 机器学习: 使用机器学习方法, 例如线性回归和SVM分类。

2 本周计划

- 继续看相关论文;
- 完成论文综述;
- 复现部分论文中的方法;

表 1: 学习方法维度统计结果

方法类型		论文
监督学习	分类	Evaluating Multiple Aspects of Coherence in Student Essays (2004, D.Higgins)
	回归	Modeling Prompt Adherence in Student Essays (2014, Persing and Ng)
无监督学习		Advanced Capabilities for evaluating student writing: Detecting off-topic essays without topic-specific training (2005, D.Higgins)
		Identifying off-topic student essays without topic-specific training data (2006)
		Unsupervised Modeling of Topical Relevance in L2 Learner Text (2016)
		Off-topic essay detection using short prompt texts (2010, Louis and D.Higgins)
		Sentence Similarity Measures for Fine-Grained Estimation of Topical Relevance in Learner Essays (2016)

表 2: 文本表示维度统计结果

方法类型	论文
概率分析	Identifying off-topic student essays without topic-specific training data (2006)
	Advanced Capabilities for evaluating student writing: Detecting off-topic essays without topic-specific training (2005, D.Higgins)
	Off-topic essay detection using short prompt texts (2010, Louis and D.Higgins)
	Unsupervised Modeling of Topical Relevance in L2 Learner Text (2016)
	Sentence Similarity Measures for Fine-Grained Estimation of Topical Relevance in Learner Essays (2016)
语义分析	Off-topic essay detection using short prompt texts (2010, Louis and D.Higgins)
	Modeling Prompt Adherence in Student Essays (2014, Persing and Ng)
	Sentence Similarity Measures for Fine-Grained Estimation of Topical Relevance in Learner Essays (2016)
机器学习	Identifying off-topic student essays without topic-specific training data (2006)
	Modeling Prompt Adherence in Student Essays (2014, Persing and Ng)