



基于社交文本的自杀倾向原因识别

刘德喜, 刘仁强, 付洪, 万常选, 刘喜平, 廖国琼
江西财经大学信息管理学院



论文摘要

自杀问题是严重的心理健康问题, 关系到个人健康发展和社会和谐稳定, 而提前发现自杀倾向原因对自杀诊断和预防自杀起到关键的作用。该文利用条件随机场CRF和基于长短期记忆网络的EAP-Bi-LSTM-CRF方法, 从具有自杀倾向的社交文本中识别自杀倾向原因。CRF模型依赖手动构建特征, 包括: 词、词性、依存句法(依存关系、到支配词距离)、自杀词典(自杀词典词、词典依存路径)。EAP-Bi-LSTM-CRF模型在基准模型Bi-LSTM-CRF基础上, 使用根据自杀数据集训练的词向量模型ELMo词向量、同时加入词性表示和注意力机制。结合识别自杀倾向原因后的应用场景, 使用四种评价方式, 分别为EM、FM、PRM、WRL。在两种方法的识别结果中, CRF在使用所有候选特征后, 在EM、FM、PRM、WRL四种评价方式的F值分别为0.586、0.629、0.683和0.757。EAP-Bi-LSTM-CRF模型在四种评价方式上的F值分别为0.621、0.677、0.7、0.801, 比基准模型Bi-LSTM-CRF分别提高了2.8、2.5、1.9、0.9个百分点, 比使用所有特征的CRF提高了3.5、4.8、1.7、4.4个百分点。

论文主要贡献

1. 构建识别社交文本中自杀倾向原因的模型;
2. 在条件随机场识别模型中, 除通用的特征外, 还结合中科院自杀词典和句法特征, 加入自杀词典词和词典依存路径特征。
3. 提出EAP-Bi-LSTM-CRF识别模型, 该模型在Bi-LSTM-CRF模型的基础上, 使用根据自杀数据集训练的词向量模型ELMo替换Word2vec词向量表示, 并根据词性在条件随机场识别中的提升效果, 在输入中加入词性表示, 最后在双向长短期记忆网络的隐藏层输出后面加入注意力机制。
4. 构建了基于微博的评测数据, 并对所提出的方法进行了实验分析。

实验仿真

模型	评价方式	P	R	F
Bi-LSTM-CRF	EM	0.618	0.569	0.593
	FM	0.68	0.62	0.652
	PRM	0.709	0.655	0.681
	WRL	0.825	0.762	0.792
E-Bi-LSTM-CRF	EM	0.644	0.592	0.617
	FM	0.702	0.645	0.672
	PRM	0.737	0.658	0.695
	WRL	0.844	0.752	0.795
A-Bi-LSTM-CRF	EM	0.649	0.570	0.607
	FM	0.706	0.620	0.660
	PRM	0.729	0.655	0.690
	WRL	0.840	0.755	0.796
P-Bi-LSTM-CRF	EM	0.649	0.570	0.600
	FM	0.706	0.62	0.656
	PRM	0.729	0.655	0.690
	WRL	0.840	0.755	0.795
EAP-Bi-LSTM-CRF	EM	0.652	0.593	0.621
	FM	0.711	0.646	0.677
	PRM	0.747	0.659	0.700
	WRL	0.855	0.754	0.801
CRF	EM	0.700	0.504	0.586
	FM	0.751	0.541	0.629
	PRM	0.779	0.608	0.683
	WRL	0.865	0.674	0.757

算法描述

输入: 已训练好的 ELMo 模型和配置文件, 分词带标签的训练集和验证集, 以及提取的字表、词表和词性标注和标签等文件

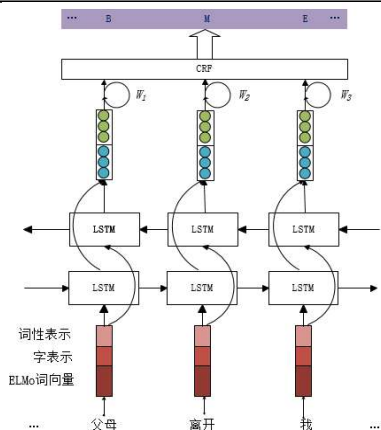
输出: 训练好的 EAP-Bi-LSTM-CRF 模型

1. 定义加载哈工大 pyltp 词性标注模块
2. 定义数据输入节点以及数据流网络结构
3. 构建注意力机制和定义 logits 函数
4. 根据 crf 定义最大熵损失函数 log_likelihood
5. 定义 adam 优化算法和包含 learning rate 等参数的训练模块
6. 定义数据初始化和保存及加载模型的对象
7. 数据转换处理(转换成关于文件中位置的 int 数值表示)
8. 模型训练


```

      for 训练数据中的每个批次(batch) {
        for 每个批次中的每个句子(s) {
          解码成中文句子后输入 ELMo 模型获得 word_representation
          使用 Bi-LSTM 获得句子中词的字向量表示 char_representation
          使用 Bi-LSTM 获得句子中词的词性表示 pos_representation
          将词向量、字向量和词性表示做 concat 拼接操作}
          利用 Bi-LSTM 获得拼接后的前向和后向输出信息
          将前后项输出信息拼接并引入注意力机制
          经 logits 函数后得到句子中词关于标签的概率矩阵
          根据 logits 输出计算损失函数值}
          训练一轮后获得验证集在当前训练模型上的精确率和 F 值并保存模型
        }
      }
      9. 重复步骤 8 且最多重复训练 nepochs 次
      if 验证集在本次训练获得的模型上的 F 值比前一轮高
        保存本次训练的模型并更新 best-F 值
      else 三轮后的 F 值没有提高, 则中断退出训练
      
```

系统模型



论文结论

- EAP-Bi-LSTM-CRF相比基准模型Bi-LSTM-CRF在EM、FM、PRM和WRL四个评价方式上的F值分别提高了2.8%、2.5%、1.9%和0.9%, 提高效果显著; 相比CRF模型分别提高了3.5%、4.8%、1.7%和4.4%。
- 动态词向量ELMo比静态的Word2vec词向量更能区分词在不同语境中的作用, 能更好地理解词与词的语义结构。
- 注意力机制调整了当前词对输出标签的贡献, 使得添加注意力机制的双向长短期记忆网络比没有添加注意力机制的双向长短期记忆网络能更好地捕获不同词的重要性。
- 受启发于CRF模型的特征工程, 在EAP-Bi-LSTM-CRF中使用词性对词的语义和语境理解具有辅助作用。
- 从社交文本中自动抽取自杀倾向原因有重要的社会意义, 但抽取效果还有待进一步提升。

