

汽车行业用户观点主题及情感识别

Just a test

nlp@cs.nju

2018 年 11 月 29 日

任务介绍

数据背景：用户在汽车论坛中对汽车相关内容的讨论或评价。

任务目标：

- 根据用户的评论文本识别出用户讨论的主题和对各主题的情感极性。
- 主题被分为10类，包括：动力、价格、内饰、配置、安全性、外观、操控、油耗、空间、舒适性。
- 情感分为3类，分别用数字0、1、-1表示中立、正向、负向。

数据集：

- 训练集：12572条。五个字段：content_id, content, subject, sentiment_value, sentiment_word
- 测试集：2753条句子。只给了：content_id, content，要求预测subject, sentiment_value。

评价指标：以pair为单位的F-1 score.

任务解析

我们采用一个pipeline的方式将该任务分解为两个分类任务，即先预测主题，再根据主题预测情感极性。

- 主题分类：

- ▶ 多标签的文本分类任务
- ▶ 我们采用Binary Relevance策略将多标签转为二分类问题。
- ▶ 在预测的时候率利用threshold来判断该主题是否存在。

- 情感分类：

- ▶ 基于角度的情感分类任务(ABSA)，输入为句子和主题的分分类问题。
- ▶ 对于同一句话不同的主题可能情感极性不同。
- ▶ 采用神经网络并利用下面两种机制来捕捉句子和主题的关系：
 - ★ attention机制
 - ★ 门控机制。

主题分类

- 卷积神经网络(CNN): [Kim, 2014]

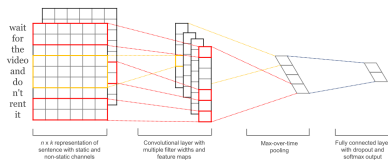
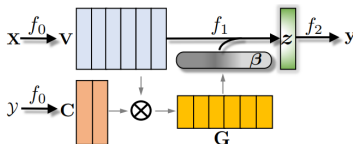


Figure 1: Model architecture with two channels for an example sentence.

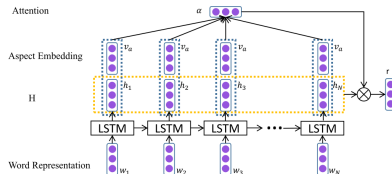
- Multi-Label Embedding Attentive Model:



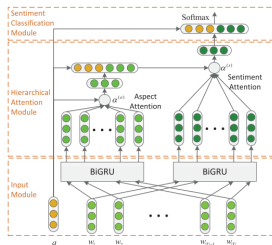
- ▶ 联合学习label的embedding，用label embedding来做attention，直接学习到和输出相关的词。
- ▶ 对于每个label有一个独立的attention过程，从而学习到各个label对应的句子表示，适合于多标签分类。

情感分类

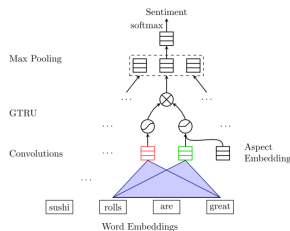
Attention_LSTM:[Wang et al., 2016]



HEAT:[Cheng et al., 2017]



GCAE:[Xue and Li, 2018]



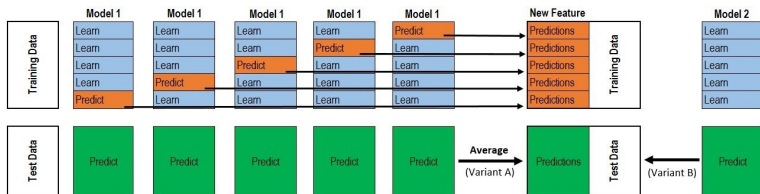
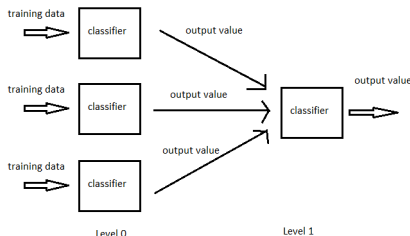
多种词向量

- Chinese-Word-Vectors
(<https://github.com/Embedding/Chinese-Word-Vectors>)
- fasttext (<https://fasttext.cc/docs/en/crawl-vectors.html>)
- Tencent AI Lab Embedding Corpus
(<https://ai.tencent.com/ailab/nlp/embedding.html>)
- 中文elmo, 来自于[ELMoForManyLangs](<https://github.com/HIT-SCIR/ELMoForManyLangs>)

模型融合

- 投票法、平均法
- Stacking: 第二层模型采用Logistic Regression.

Concept Diagram of Stacking



BERT

我们最后尝试了一下谷歌的BERT，在谷歌预训练好的中文BERT上做微调[Kenton et al., 2017]

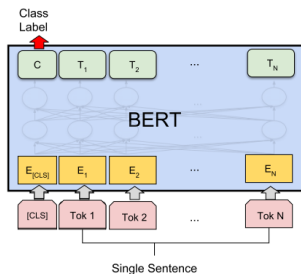


图: 主题分类

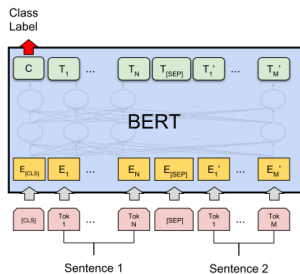
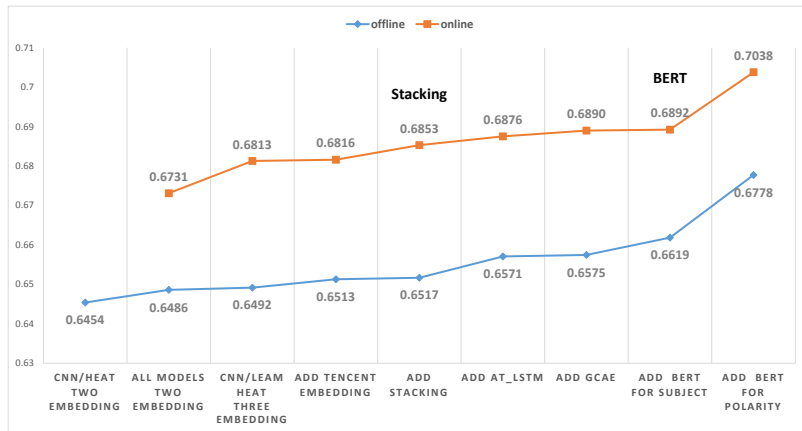


图: 情感分类

BERT非常适合于做ABSA, 因为它可以接受pair的输入, 而且是由基于self-attention的transformer构成。

模型改进折线图：线下为从训练集中分出的测试集。



失败的尝试

- 数据不平衡
- 其他的多标签学习策略
- 非深度学习方法: xgboost, LightGBM
- 将第二层LR换成xgboost等复杂模型会导致过拟合
- 三层stacking
- 伪标签
- 失败的差异性: 分词, 随机种子
- 搜索threshold
- 训练领域相关的词向量
- More....

未来可以改进的

- 利用标注好的情感词信息。
- BERT的领域迁移
- 从实用角度来说，很多模型还是太大了，所以可以考虑适当降低精度，压缩模型。
- 联合学习

References I



Cheng, J., Zhao, S., Zhang, J., King, I., Zhang, X., and Wang, H. (2017). Aspect-level Sentiment Classification with HEAT (HiErarchical ATtention) Network. *Proceedings of the 26th ACM International Conference on Information and Knowledge Management*, pages 97–106.



Kenton, M.-w. C., Kristina, L., and Devlin, J. (2017). BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding.



Kim, Y. (2014). Convolutional Neural Networks for Sentence Classification. pages 1746–1751.



Wang, Y., Huang, M., Zhao, L., and Zhu, X. (2016). Attention-based LSTM for Aspect-level Sentiment Classification. *Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pages 606–615.



Xue, W. and Li, T. (2018). Aspect Based Sentiment Analysis with Gated Convolutional Networks.