**Sentiment Analysis**

**Introduction**

情感分析是自然语言处理领域的热点研究问题，主要思想是通过对自然语句进行一系列处理，获取该语句的情感倾向，挖掘深层信息。在互联网已经广泛普及的现代社会，每天出现在互联网上的评论信息数以亿计，如果能对如此庞大的信息数据集加以有效利用，将产生很高的社会效益。情感分析通常在词的维度上进行。首先需要对相关文本进行分词处理，然后根据已有的情感字典，对分好的词语经过相关模型的一系列处理，得到一个有关该文本的情感分析结果。目前经常用来进行情感分析的模型有Naive Bayes、SVM、Logistic Regression等。

**Project Task**

* 针对网络上已有的语料信息，使用经典结构模型，对语料进行情感分析，使得训练后的模型能够有效对新输入的语料进行情感的判别。
* 尝试对经典模型进行优化，使结果更准确。鼓励使用deep learning相关结构来考虑。

**Schedule**

第2周：中秋节放假

第3-4周：搭建所需环境python+TensorFlow，推荐使用pyCharm作为集成开发环境，了解python和TensorFlow框架的使用。

第5周：国庆节放假

第6-8周：尝试搭建经典模型，选定一个数据集进行训练，并给出训练后的测试结果。

第9-10周：对多个经典模型(至少3个)进行评估，给出评估指标，作为今后优化的依据。

第11-13周：尝试对模型进行优化，或者设计自己的模型（推荐deep learning相关知识）

第14-15周：结果汇总，制作报告，撰写论文

第16-17周：presentation

**Resources：**

* 简易版教程与demo：https://github.com/aditya-xq/Text-Emotion-Detection-Using-NLP
* 数据集：  
  微博情感分析测评数据：https://pan.baidu.com/s/1psjysSXpKOEb1ciem7DsRw 密码：7hb4  
  中文对话情绪语料：https://github.com/xxxspy/Chinese\_conversation\_sentiment
* 相关文献：  
  文本情感分类：<https://pdfs.semanticscholar.org/c804/78e361ed8f5fd5400fdbd4f6a6f37a2e4b57.pdf>  
  deep learning与情感分析：<https://link.springer.com/content/pdf/10.1007%2Fs13042-018-0799-4.pdf>
* 安装教程：

Anaconda3(python环境管理器)安装教程：<https://www.jianshu.com/p/026a2c43b081>

pyCharm安装及配置教程：<https://www.runoob.com/w3cnote/pycharm-windows-install.html>