**一、代码结构**

├── sentiment.py # 情感倾向分析主函数，包括训练、预估、预测部分

├── nets.py # 本例中涉及的各种网络结构均定义在此文件中

├── utils.py # 定义通用的函数，例如加载词典，读入数据等

├── C-API #模型预测C-API接口

**二、简介**

情感倾向分析针对带有主观描述的中文文本，可自动判断该文本的情感极性类别并给出相应的置信度。情感类型分为积极、消极、中性。

**三、模型概览**

nets.py 中包含一下模型：

bow\_net：Bow(Bag Of Words)模型，是一个非序列模型。使用基本的全连接结构。

cnn\_net：浅层CNN模型，是一个基础的序列模型，能够处理变长的序列输入，提取一个局部区域之内的特征。

gru\_net：单层GRU模型，序列模型，能够较好地解序列文本中长距离依赖的问题。

lstm\_net：单层LSTM模型，序列模型，能够较好地解决序列文本中长距离依赖的问题。

bilstm\_net：双向单层LSTM模型，序列模型，通过采用双向lstm结构，更好地捕获句子中的语义特征。AI平台上情感倾向分析模块采用此模型进行训练和预测。

**四、数据准备**

数据格式：每一行为一条样本，以`\t`分隔，第一列是类别标签，第二列是输入文本的内容，文本内容中的词语以空格间隔。以下是两条示例数据：

```

2 特 喜欢 这种 好看的 狗狗

2 这 真是 惊艳 世界 的 中国 黑科技

1. 环境 特别 差 ，脏兮兮 的 ，再也 不去 了

代码中采用jieba对句子进行分词处理后，再作为模型的输入。

**五、模型训练与预测**

**1.模型训练**

python sentiment\_classify.py \

--train\_data\_path ./data/train\_data/corpus.train \ # 训练数据路径

--word\_dict\_path ./data/train.vocab \ # 词典路径

--mode train \ # train模式

--model\_path ./models # 模型保存路径

**2.模型评价**

python sentiment\_classify.py \

--test\_data\_path ./data/test\_data/corpus.test \# 测试数据路径

--word\_dict\_path ./data/train.vocab \ # 词典路径

--mode eval \ # eval模式

--model\_path ./models/epoch0/# 预测模型路径

**3.模型预测**

python sentiment\_classify.py \

--test\_data\_path ./data/test\_data/corpus.test \# 测试数据路径

--word\_dict\_path ./data/train.vocab \ # 词典路径

--mode infer \ # infer模式

--model\_path ./models/epoch0/# 预测模型路径