**基于周杰伦歌词的数据挖掘**

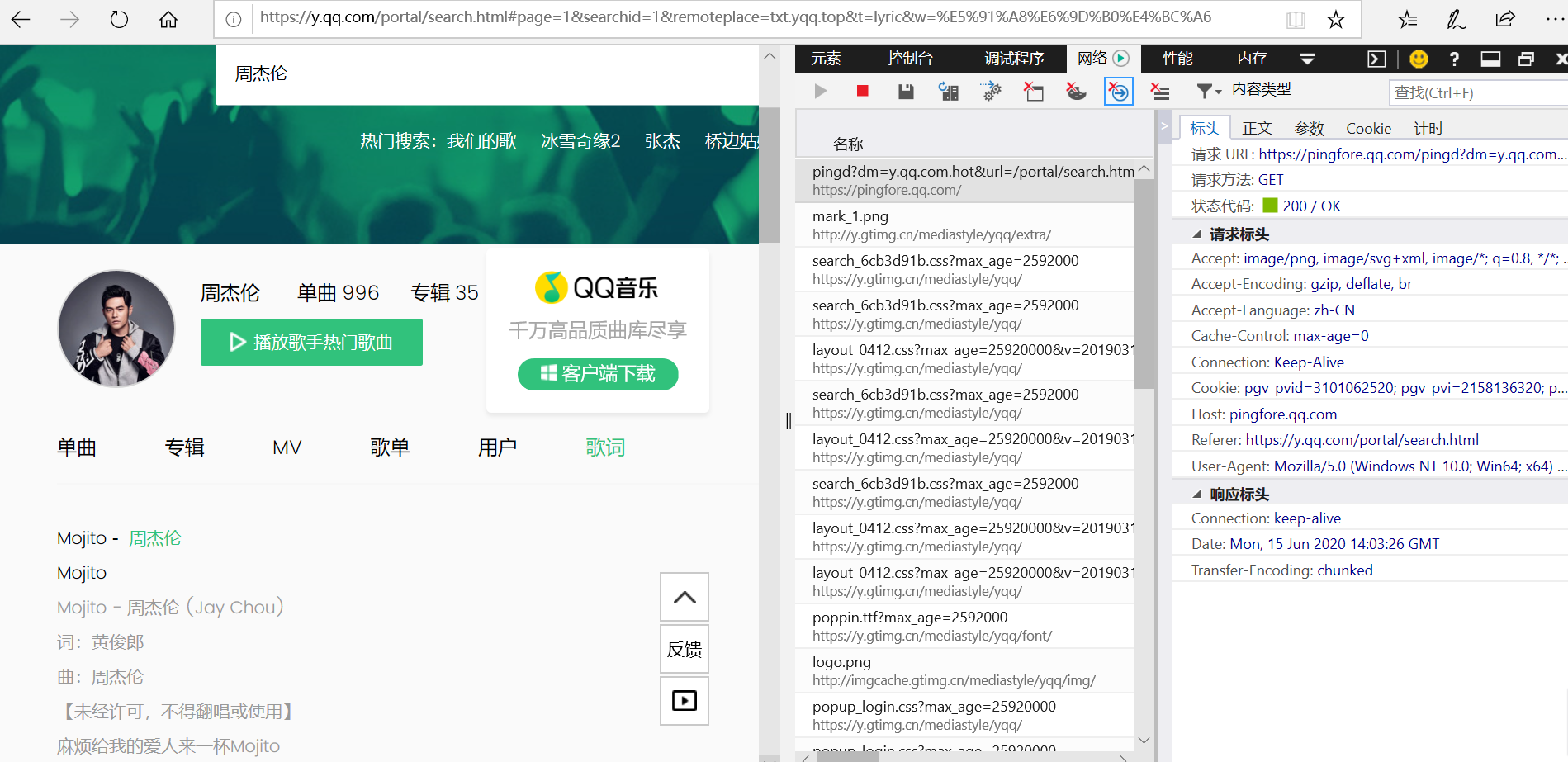
**Task 1 收集相关数据，可视化周杰伦歌词用词习惯；**

**Task 2 对歌词进行情感分析，构建歌词情感预测模型；**

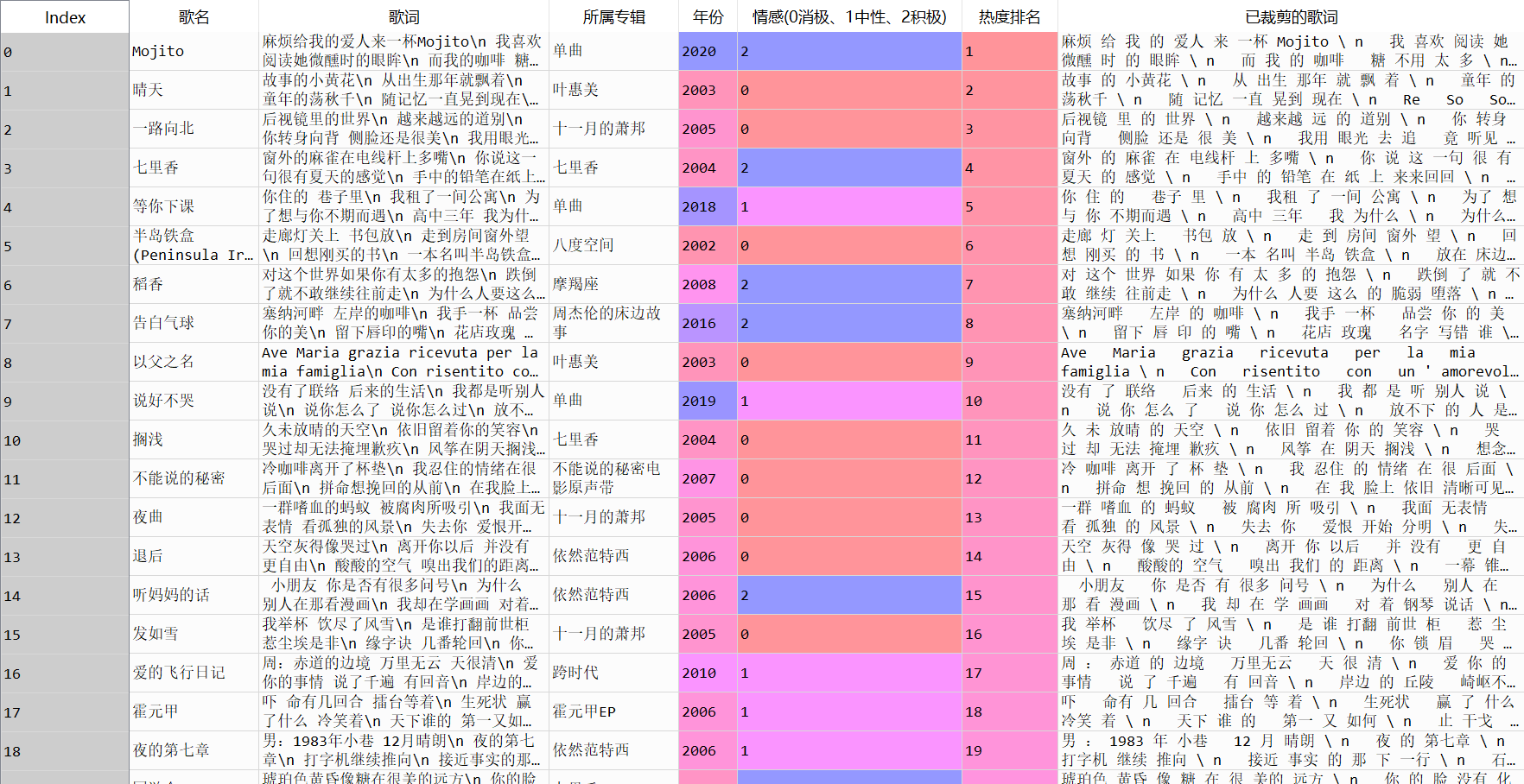
**Task 3 构建“周氏词风”生成器**

**第一部分 数据获取与预处理**

1）Python爬虫技术获取数据（QQ音乐中周杰伦歌曲近期热度排行榜top60的歌曲信息）

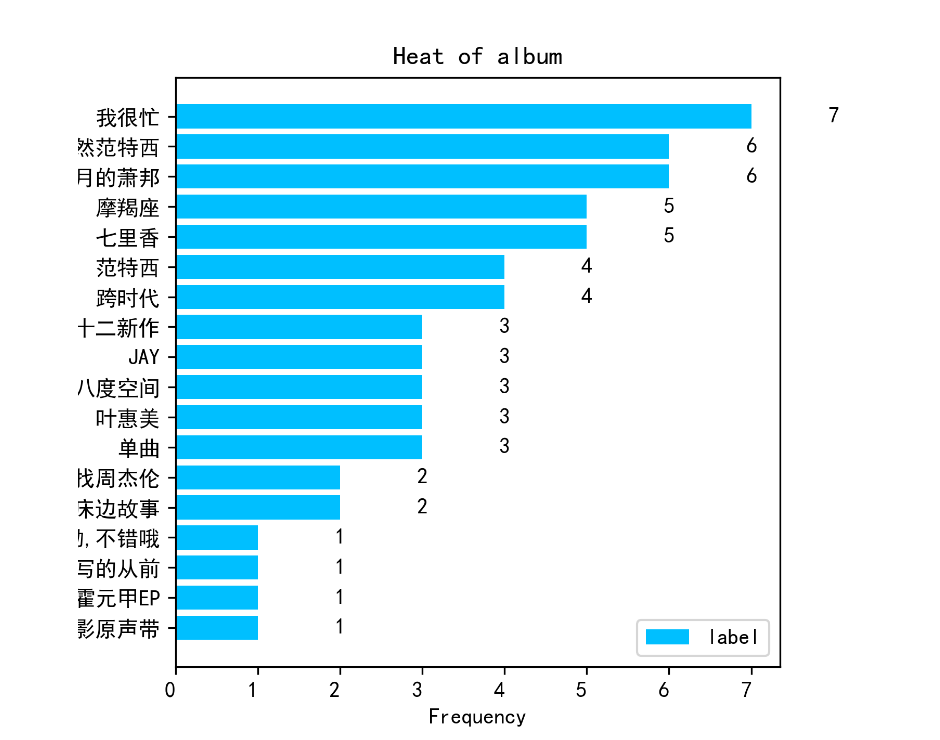


2）解析数据、数据标注后利用jieba将歌词分词

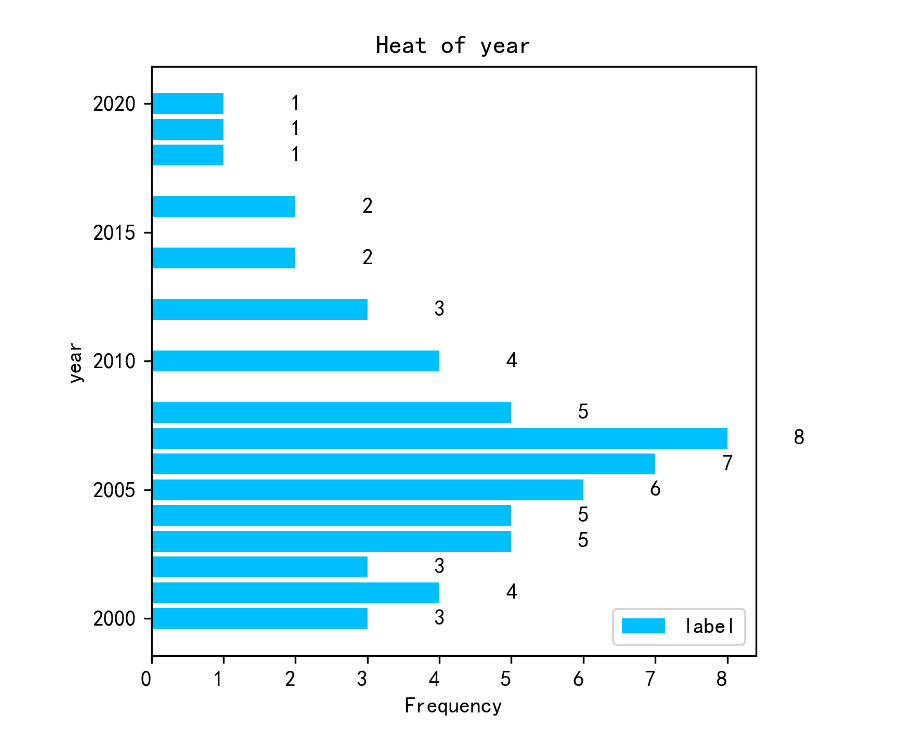


**第二部分 描述性统计分析**

专辑热度



年份热度



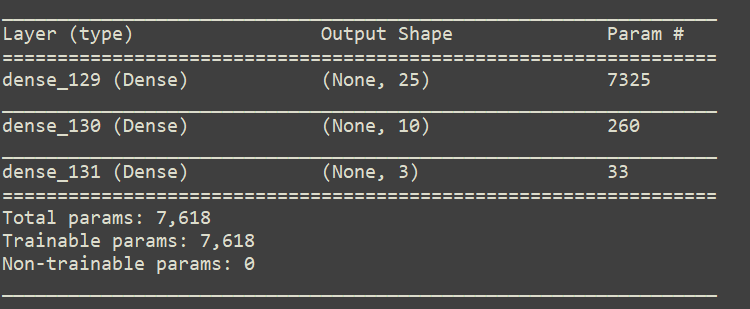
wordcloud词频统计并可视化词云图：



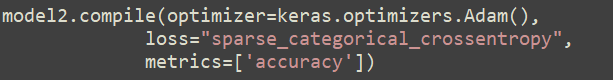
**第三部分 情感预测**

1. **基于snow库方法的情感分析（尚未完成）**
2. **深度学习方法**

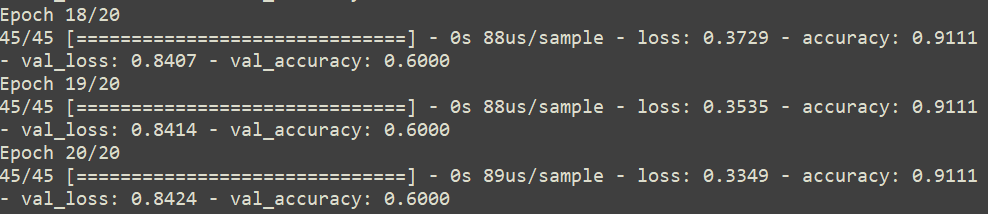
1）建模 双隐层DNN



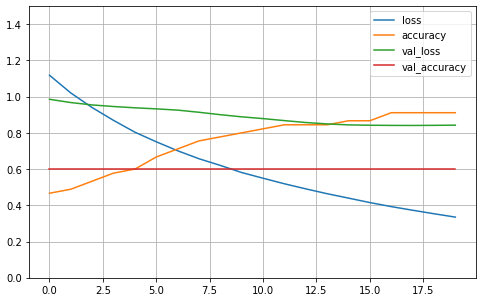
2）设置Loss与优化方法



3）训练过程展示

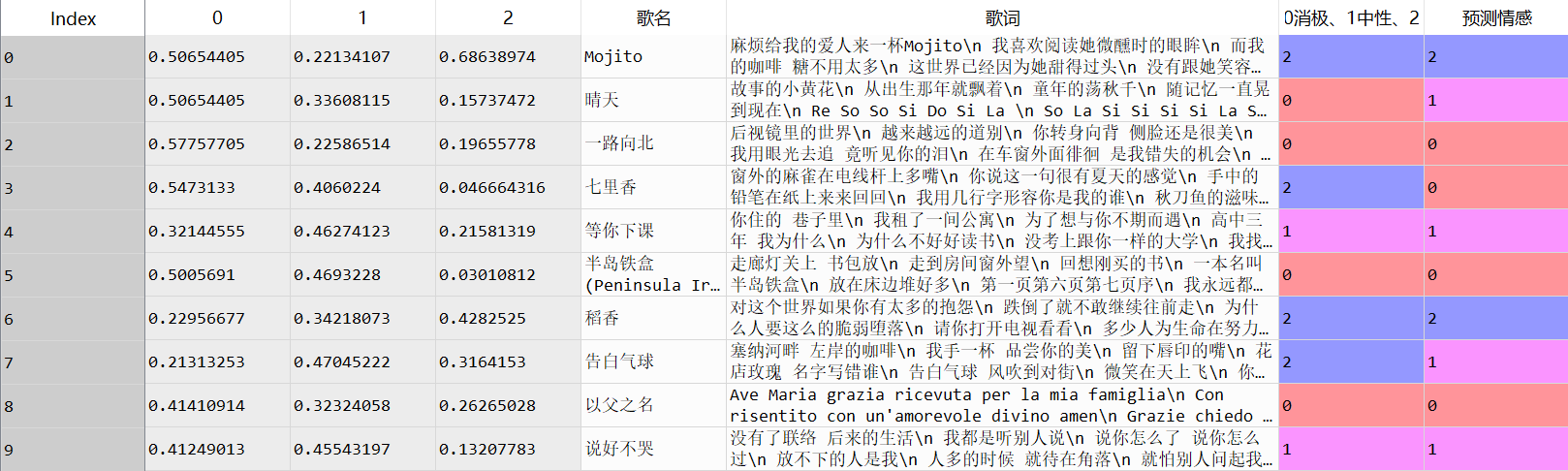


4）可视化学习曲线



5）预测





进阶：考虑上word2vec模型 或 RNN（尚未完成）

**第四部分 歌词生成器（尚未完成）**