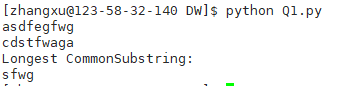
题目一：计算最长公共子序列

An：

代码见附件Q1.py；

运行截图如下：



题目二：完成self-attention计算过程

An：

计算过程如下：

其中，，，，表示输入数据，分别为query/key/value对应可训练权重；

题目三：将给定的数据表格进行数据结果统计，并按要求格式输出

An：

代码见附件Q3.py；

输出数据表格见Q3.json；

思路：

最终输出是以 “系统编号”为identity的，并计算不同算法下的不同指标；在原始数据中，不同编号中包含有不同的句子，而每个句子可能被不同算法召回，故构造了 系统编号🡪句子🡪算法🡪结果 的树用于统计计算。

其中，稍有疑问的是对于查全率的统计，由于题目并未详细说明计算预测错误或正确的查全率，本结果中计算各算法计算正确的查全率。

题目四：在不依赖其他框架的基础下，实现MLP模型及训练过程

An：

代码见Q4.py；

题目五：分类模型训练评估

An:

代码见Q5.py。