

# 基础理论部分

## 0. Can you come up out 3 sceneraies which use AI methods?

Ans: 语义搜索, 智能问答, 扫地机器人

## 1. How do we use Github; Why do we use Jupyter and Pycharm;

Ans: 使用Github进行项目代码管理、开源与协作, 同时也学习github上的优秀开源项目; 使用jupyter进行课堂编程演示; 使用Pycharm进行线下项目编程

## 2. What's the Probability Model?

Ans:描述不同变量之间概率关系的模型

## 3. Can you came up with some sceneries at which we could use Probability Model?

Ans:给定一些条件预测股市涨跌的概率; 根据某些症状判断患上某种疾病的概率

## 4. Why do we use probability and what's the difficult points for programming based on parsing and pattern match?

Ans:我们之所以使用概率是因为它给我们提供了量化不确定性的方法, 使我们能够作出不确定性的声明, 并在不确定的情况下进行推理; 概率告诉我们AI系统如何推理, 并使我们能够从理论上分析AI系统的行为。

使用parsing and pattern match的难点在于, 我们无法穷尽所有的pattern。

## 5. What's the Language Model;

Ans:语言模型是指一个单词序列上的概率分布。

## 6. Can you came up with some sceneraies at which we could use Language Model?

Ans:语音识别，机器翻译，词性标注，句法分析等

### **7. What's the 1-gram language model;**

Ans:指不考虑当前词的上下文环境，单个词出现的概率，一个句子出现的概率等于句子中每个单词概率的乘积

### **8. What's the disadvantages and advantages of 1-gram language model;**

Ans:优点容易计算，缺点是不符合实际情况，句子中的单词实际上存在相互关系

### **9. What't the 2-gram models;**

Ans:指当前词的概率与其前一个词有关，是一种上下文有关的模型

编程实践部分

Q: 这个模型有什么问题？ 你准备如何提升？

Ans:使用的训练语料集中在某一领域，可能导致训练的模型有偏。引入更丰富的训练语料；