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

Ans: The recommendation systems, the voice recognition techniques, the data analytics using machine/deep learning algorithms

1. How do we use Github? Why do we use Jupyter and Pycharm?

Ans: You put all course related materials in Github so that we can check all the time. And we just upload our assignments into Github. The reason why we use Pycharm is that it provides smart code editor which supports multiple languages and it has a huge collection of tools out of the box, including an integrated debugger and a test runner, which is relatively convenient for us to use.

2. What's the Probability Model?

Ans: A probability model is a mathematical representation of a random phenomenon. It is defined by its sample space, events within the sample space, and probabilities associated with each event.

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

Ans: In some searching engines, when we input some key words, the engine will use the probability model to rank the relevant documents to reply to our requests. Sometimes, we use the probability model to make predictions on future trends of certain things, like stock/housing prices.

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

Ans: Probability is a direct measure of relevance I think. Large probability means high frequency of using something or close relevance of the results to our requests. Pattern matching is the act of checking a given sequence of tokens for the presence of the constituents of some pattern. In contrast to pattern recognition, the match usually has to be exact: "either it will or will not be a match". So I guess this is one of the biggest challenge for implementing the pattern matching compared with the pattern recognition as it requires accuracy. Parsing algorithms often rely on pattern matching to transform strings into syntax trees. I guess sometimes the tree structures are very complex or strings are hard to transform due to complexity issues. What’s more, the data sparsity could also be a problem.

5. What's the Language Model?

Ans: The Language Model is a subset of the probability model. It measures the frequency of using certain sentences or words in a large context. The input is sentences/words and the output is the probability between 0 to 1.

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

Ans: Again, when we input some key words in the searching engine, the engine will help us complete the searching requests by using the language model to recommend the rest part of our search requests. In the speech recognition, we also use the language model to recognize what people say according to the probability.

7. What's the 1-gram language model?

Ans: The formula is given as $Pr(W_i) = \frac{Count(Wi)}{N}$. It means that the probability of certain word in a large context is calculated as the ratio of the number of that word over the total number of words.

8. What's the disadvantages and advantages of 1-gram language model?

Ans: The advantage of the 1-gram language model is that it can get the probability of certain words easily so that we can have a direct feeling of how the words are used in the context. When we just care about the usage of certain words, the 1-gram language model is a good choice. But most of the time, we care much more than one word, then the 1-gram language model does not work due to the limitations. So 1-gram language has a narrow range. It only deals with simple words.

9. What's the 2-gram models?

Ans: The formula is given as $Pr(W_{i+1}|W_i) = \frac{Count(Wi,Wi+1)}{Count(Wi)}$, which measures the probability of using the two adjacent words or sequences in a large context.