Question list:

1. What are three parts of a search engine? How does a page-ranking algorithm work?

1、搜索引擎的三个部分是什么？页面排序算法是如何工作的？

搜索引擎的三部分是：spider、index 和搜索引擎软件。页面排序算法是按关键词的位置和频率排序的。频率高的排在前面。

The three parts of the search engine are: spider, index, and search engine software. The page sorting algorithm is sorted by the location and frequency of the key words. The high frequency is in the front.One elements a search engine algorithm scans for is the frequency and location of keywords. Those with higher frequency are typically considered more relevant.

2. Give at least 5 reasons that make silicon valley so successful.

2、给出至少 5 个使硅谷如此成功的理由。

硅谷成功的理由：（1）硅谷有世界上最优秀的计算机的专业人才和最好的支持服务，还有有名的大学。（2）在硅谷，人们不害怕失败而且敢于冒险。（3）硅谷有大量的风险投资人。（4）年轻人喜欢创业胜过薪酬。（5）硅谷的人都很勤奋。

Silicon Valley's reasons for success: (1) Silicon Valley has the world's best computer professionals and the best support services, as well as famous universities. (2) in Silicon Valley, people are not afraid of failure and dare to take risks.

（3）There are a lot of venture capitalists in Silicon Valley.

（4） young people like to start a business better than pay.

（5）The Valley's professionals are among the most hardworking people anywhere.

3. Suppose you are a interviewer, give some of your advice in terms how to interview an architect.

3、假设你是一位面试官，就如何采访一位建筑师给出一些建议。

4. In the field of Information Retrieval, TF-IDF theory is basic one. Try to 1) explain it 2)how it can be fitted into the idea of Vector-Space-Model?

4、在信息检索领域，TF-IDF 理论基础。尝试 1）解释它 2）它如何能被装配到向量空间模型的想法中？

TF-IDF（term frequency–inverse document frequency）是一种用于信息检索与数据挖掘的常用加权技术。TF 意思是词频(Term Frequency)，IDF 意思是逆向文件频率(Inverse Document Frequency)。

TF-IDF is a commonly used weighting technique for information retrieval and data mining. TF indicates the frequency of the entries appearing in the document D.If the fewer documents containing the word t, the smaller the N, the larger the IDF, the word t has a good classification ability.

5. What is phishing? Use an example to illustrate how a fraudulent/malicious website work?

5、什么是网络钓鱼？使用一个例子来说明欺诈/恶意网站是如何工作的？

钓鱼网站伪装成一个合法的银行或信用卡公司，用户被邀请提供他们的识别信息。这类网站是专门用来窃取用户的密码、PIN 号码，社会安全号码和其他保密信息。

Phishing website is disguised to look like a legitimate one — of a bank or a credit card company, and users are invited to provide their identifying information. Sites of this kind are used solely to steal users' passwords, PIN numbers, SSNs and other confidential information.

问题是，该木马包含一个键盘记录程序，潜伏在后台到被感染的电脑用户访问一个指定的网站。然后键盘记录来生活做了用于窃取信息。

The problem is that this Trojan contains a keylogger which lurks at the background until the user of the infected PC visits one of the specified websites. Then the keylogger comes to life to do what it was created for — to steal information.

6. Speaking of machine learning, what is the definition? The history of ML? The types of ML?

6、说到机器学习，定义是什么？ML 的历史？ML 的类型？

机器学习是用数据或以往的经验，以此优化计算机程序的性能标准。

Machine learning is programming computers to optimize a performance criterion using example data or past experience.

第一阶段是在 20 世纪 50 年代中叶到 60 年代中叶，属于热烈时期。

第二阶段是在 20 世纪 60 年代中叶至 70 年代中叶，被称为机器学习的冷静时期。

第三阶段是从 20 世纪 70 年代中叶至 80 年代中叶，称为复兴时期。

The first stage is in the middle of the 1950s to the middle of 60s, which belongs to the warm period.

The second stage, in the middle of the 1960s, to the middle of the 70s, is called the cool period of machine learning.

The third stage, from the middle of the 1970s to the middle of the 80s, is called the rejuvenation period.

Machine learning can be divided into supervised learning and unsupervised learning.

监督学习(supervised learning)

非监督学习(unsupervised learning)

7. Please spell out the brief methodological history of NLP. And name two examples that's considered to be AI-complete.
7、请简述 NLP 的简要方法学历史。并列举两个被认为是 AI 完成的示例。
NLP 成功的例子有：聊天机器人和机器翻译。
NLP's successful examples are chat robots and Machine Translation.

8. Name some common techniques that are employed to enhance network security.
8、列举一些用于提高网络安全性的常用技术。

用于提高网络安全性的常用技术：
The common technologies used to improve network security are: secret technology, digital certificates and firewalls.
用备份技术来提高数据恢复时的完整性。
Use backup technology to improve the integrity of data recovery.
审计系统
Audition system
加密
Using encryption technology

9. What is VR? What a HMD is used for?
9、VR 是什么的？什么是头盔显示器是用来做什么的？

虚拟现实技术是一种可以创建和体验虚拟世界的计算机仿真系统，它利用计算机生成一种模拟环境，是一种多源信息融合的、交互式的三维动态视景和实体行为的系统仿真使用户沉浸到该环境中。

Virtual reality technology is a computer simulation system can create and experience the virtual world, it is the use of computer to generate a simulation environment, a simulation system of 3D dynamic interactive multi-source information fusion, the visual entity behavior and users immerse in the environment.

HMD 通常采取的形式头安装在眼睛前方的屏幕护目镜。一些模拟包括额外的感官信息，并通过扬声器或耳机提供声音。
HMDs typically take the form of head-mounted goggles with a screen in front of the eyes. Some simulations include additional sensory information and provide sounds through speakers or headphones.

10. Give a definition of Software Engineering and at least four sub-areas of Software Engineering.
10、给出软件工程的定义和至少四个软件工程的子领域。

软件工程是一门研究用工程化方法构建和维护有效的、实用的和高质量的软件的学科。

Software engineering is a subject that studies the construction and maintenance of effective, practical and high quality software with engineering methods.
软件测试、软件分析、软件质量评测、软件设计
Software testing, Software analysis, Software quality evaluation, software design

11. What's the resent break through in the area of Machine Learning?
11、机器学习领域的最新突破是什么？

阿尔法围棋

AlphaGoZero 的能力则在这个基础上有了质的提升。最大的区别是，它不再需要人类数据。也就是说，它一开始就没有接触过人类棋谱。研发团队只是让它自由随意地在棋盘上下棋，然后进行自我博弈。

The ability of AlphaGoZero has been improved on a previous basis. The biggest difference is that it no longer needs human data. That is to say, it has no contact with human chess. The R & D team only lets it play chess freely and freely on the chessboard, and then plays the game of self.

12. To explain CIA triad in the area of information security(InfoSec).
12、解释在信息安全领域的情报局黑社会（信息安全）。

13. What is denial-of-service attack (DoS attack)? How does DDoS work?
13、拒绝服务攻击（DoS 攻击）是什么？DDoS 工作怎么样？

即是利用网络上已被攻陷的电脑作为“僵尸”，向某一特定的目标电脑发动密集式的“拒绝服务”式攻击，用以把目标电脑的网络资源及系统资源耗尽，使之无法向真正正常请求的用户提供服务。

DOS is the use of computer on the network has been compromised as a "zombie", launched intensive to a specific target computer "denial of service" attacks,

the cyber source and target computer system resource depletion, which can not provide services to the real normal request user.

DDoS 攻击可以具体分成两种形式：带宽消耗型以及资源消耗型。它们都是透过大量合法或伪造的请求占用大量网络以及器材资源，以达到瘫痪网络以及系统的目的。

DDoS attacks can be divided into two forms: bandwidth consumption and resource consumption. They all occupy a large number of network and equipment resources through a large number of legitimate or forged requests to achieve the purpose of the paralyzed network and the system.

14. Use 150 words each to tell us about these six person: Jack Ma, Robbin Li, Feifei Li, Geoffrey Hinton, Yann LeCun, and Mark Zurkerberg.
14、用 150 字告诉我们这六个人：Jack Ma，Robbin Li，李菲菲，Geoffrey Hinton，Yann LeCun，Mark Zurkerberg。

Yann LeCun (/ləˈkʌn/; born 1960) is a computer scientist with contributions in machine learning, computer vision, mobile robotics and computational neuroscience.
He is well known for his work on optical character recognition and computer vision using convolutional neural networks (CNN), and is a founding father of convolutional nets. He is also one of the main creators of the DjVu image compression technology.
Yann LeCun （/ləˈkʌn /； 生于 1960）是一个在 机器学习 贡献的计算机科学家，计算机视觉，移动机器人 和 计算神经科学。他是众所周知的，他的工作 光学字符识别 和 使用 卷积神经网络 计算机视觉 （美国有线电视新闻网），是卷积网创始人。 他也是其中的 DjVu 图像压缩技术的主要创造者。
Mark Elliot Zuckerberg (/ˈzʌkərbɜːrg/; born May 14, 1984) is an American computer programmer and Internet entrepreneur. He is a co-founder of Facebook, and currently operates as its chairman and chief executive officer. Zuckerberg launched Facebook from his Harvard University dormitory room on February 4, 2004. Since 2010, Time magazine has named Zuckerberg among the 100 wealthiest and most influential people in the world as a part of its Person of the Year award. In December 2016, Zuckerberg was ranked 10th on Forbes list of The World's Most Powerful People.
埃利奥特扎克伯格 （/ˈZʌKəRBɜːRg/；生于 1984 年 5 月 14 日）是美国的一个计算机程序员和互联网企业家。他是一个 脸谱网创始人之一，目前经营的董事长兼首席执行官。 扎克伯格推出了脸谱网从他的 哈佛大学 宿舍 2004 年 2 月 4 日。自 2010 以来， 时间 杂志评选扎克伯格在 100 最富有和最有影响力的人在世界上作为其 年度人物 奖的一部分。 2016 年 12 月，扎克伯格在 福布斯 全世界最有权力人物榜排名第十。

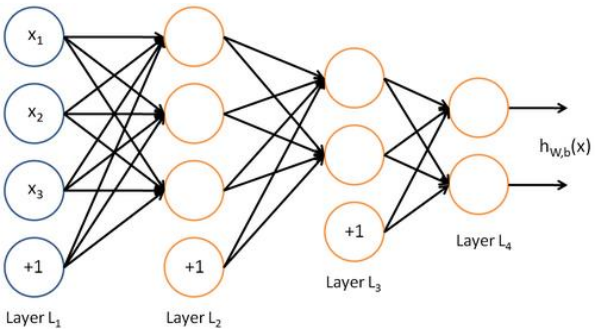
15. Why we say “Geoffrey Hinton is someone who can be hailed as guru of new computing era”?
15、为什么我们说“Geoffrey Hinton 可以被誉为大师的新计算时代”？

16. Use 150 words or so to explain the concept of Convolutional Neural Network, including its structure and its feature.

16、用 150 个字来解释卷积神经网络的概念，包括卷积神经网络的结构及其特点。

卷积神经网络（Convolutional Neural Network,CNN）是一种前馈神经网络，它的人工神经元可以响应一部分覆盖范围内的周围单元，对于大型图像处理有出色表现。[2] 它包括卷积层(convolutional layer)和池化层(pooling layer)。
Convolutional Neural Network (CNN) is a feed-forward neural network. Its artificial neurons can respond to a number of surrounding units in a range of coverage, and have excellent performance for large image processing. It includes [2] convolution (convolutional layer) layer and pool layer (pooling layer).

17. 1) Explain the Neural Network constructed by the figure below 2) What is Activation Function? Use an example to illustrate it.
17、1）解释图 2 所示的神经网络）什么是激活函数？用一个例子来说明它。



In this figure, we have used circles to also denote the inputs to the network. The circles labeled "+1" are called **bias units**, and correspond to the intercept term. The leftmost layer of the network is called the **input layer**, and the rightmost layer the **output layer** (which, in this example, has only one node). The middle layer of nodes is called the **hidden layer**, because its values are not observed in the training set. We also say that our example neural network has 3 **input units** (not counting the bias unit), 3 **hidden units**, and 1 **output unit**.

在这个图中，我们使用圆来表示对网络的输入。圆圈标记为“＋1”被称为 偏置单元和对应的截距项。网络的最左边的一层叫 输入层和输出层最右边的层 （其中，在这个例子中，只有一个节点）。节点的中间层被称为 隐藏层，因为它的值在训练集观察不到。我们还说我们的神经网络有 3 输入单元 （不包括偏置单元），3 隐藏单位，和 1 输出单元。

指如何把“激活的神经元的特征”通过函数把特征保留并映射出来（保留特征，去除一些数据中是的冗余），这是神经网络能解决非线性问题关键。

The activation function refers to how to retain and map the characteristics of the activated neuron by function, which preserves

the feature and removes the redundancy in some data, which is the key to solve the nonlinear problem by neural network.

18. What does Trojan Horse means in Network Security? How does it work to cause danger for unsuspected users?
18、特洛伊木马在网络安全中意味着什么？它怎么给使用者造成危险吗？

如果我们的计算机访问了 Trojan Horse，我们的计算机将不能正常工作并且我们信息将被泄露。
If our computer has access to Trojan Horse, our computer will not work properly and our information will be leaked.

19. Please give some guidelines in terms of News Writing.
19、请给出一些关于新闻写作的指导原则。
（1）新闻写作是倒金字塔的形式；
（2）要把重要的信息写在开头；
（3）摘要很重要一定言简意赅。
(1) news writing is the form of inverted Pyramid;
(2) to write important information at the beginning;
(3) the very important must be concise and comprehensive.

20. Do you think this course(Computer English) useful/helpful? If so, how can it be useful/helpful? According to your own opinion, what is your suggestion to this course and to the teacher in order to achieve a better learning experience and results?
20、你认为这门课（计算机英语）有用吗？如果是的话，它怎么有用/有帮助？根据你自己的意见，你对这个课程和老师有什么建议，以获得更好的学习经验和成绩？

我认为 Computer English 是有用的。它使我的视野更开阔了比以前。它让我知道了很多关于计算机领域的知识。例如：人工智能、虚拟现实、增强现实、软件工程、等等。在我看来，我建议这门课应该就一个专题较深入的讲解，在一个学期内。不然总是觉得上这门课就是去了解新名词关于计算机的。

I think Computer English is useful. It makes my view more open than before. It gives me a lot of knowledge about the computer field. For example: artificial intelligence, virtual reality, augmented reality, software engineering, and so on. In my opinion, I suggest that this course should be explained in depth on a topic, in one semester. Otherwise, we always think that this course is to understand the new nouns about computers.

I am now uploading the Question List I had mentioned last week.
The final exam(in written form and oral from) questions will be picked out from this 20-question-list. What you need to do is to get the answer to each of these 20 questions and be familiar with all these answers.

No worries, I am not a harsh grader. But you need get prepared for it.
Ni Men Jia You!
我现在上传我上周提到的问题清单。
期末考试（笔试和口试）将从这 20 个问题列表中选出。
你需要做的是得到这 20 个问题的答案，并熟悉所有这些答案。
不用担心，我不是一个严厉的平地工。但你需要为此做好准备。
倪家佳！