1. 人工智能、机器学习、深度学习的异同。

机器学习是实现人工智能的一种方法。深度学习是实现机器学习的一种技术，三者的关系如下图：



人工智能：只要模仿出人类的行为，都称之为人工智能，深度学习和机器学习都是人工智能的范畴。

机器学习：强调学习，通过对数据模式的自识别，输出正确的结果。

深度学习：机器学习的一种，包含特殊的数学模型，利用大量的数据（少量数据不理想），来预测或者输出最终结果。

以下一段话，形象的区分人工智能和机器学习：

如果有一段能够模仿人类行为的程序，这就称之为AI，但是除非它的参量都是自动从数据中学会的，否则就不能是机器学习。

机器学习与深度学习对比如下：

1. 数据依赖：深度学习需要依赖大量数据，表现才佳。
2. 硬件依赖：一般深度学习依赖高端设备，传统机器学习依赖低端设备。
3. 执行时间：深度学习需要较长的时间来训练。
4. 介绍三个以上深度学习应用实例。
5. Let there be color！：给黑白照片和视频自动上色。
6. Face2Face：实时捕获面部表情，调换到另外一个人的面部。
7. LipNet：利用深度学习的读唇程序。
8. Neural Talk：根据照片内容描述照片。
9. Skype：语音实时翻译，能够翻译成不同的语言。
10. Google mail：自动回复电子邮件。
11. 监督学习与非监督学习的异同。

监督学习：从给定的训练数据集中“学习”出一个函数，当新的数据到来时，可以根据这个函数预测结果，其训练集要求包括输入和输出、即特征和目标，需要由人类事先标注。

非监督学习：从训练数据集中的“学习”过程中，只提供特征而不提供目标值，让其自行归纳，不需要人工干预，所以也被称之为归纳性学习。

区别：

1. 有无目标值：监督学习有目标值，非监督学习没有。
2. 是否人工干预：监督学习需要，非监督学习不需要。
3. 主要用途：监督学习主要用于分类和回归，非监督学习主要用于聚类。
4. 强化学习的基本原理。

[强化学习](https://www.baidu.com/s?wd=%E5%BC%BA%E5%8C%96%E5%AD%A6%E4%B9%A0&tn=SE_PcZhidaonwhc_ngpagmjz&rsv_dl=gh_pc_zhidao" \t "/home/foreverhappy/文档\\x/_blank)是从动物学习、参数扰动[自适应控制](https://www.baidu.com/s?wd=%E8%87%AA%E9%80%82%E5%BA%94%E6%8E%A7%E5%88%B6&tn=SE_PcZhidaonwhc_ngpagmjz&rsv_dl=gh_pc_zhidao" \t "/home/foreverhappy/文档\\x/_blank)等理论发展而来，其基本原理是：  
如果Agent的某个行为策略导致环境正的奖赏(强化信号)，那么Agent以后产生这个行为策略的趋势便会加强。Agent的目标是在每个离散状态发现最优策略以使期望的折扣奖赏和最大。  
[强化学习](https://www.baidu.com/s?wd=%E5%BC%BA%E5%8C%96%E5%AD%A6%E4%B9%A0&tn=SE_PcZhidaonwhc_ngpagmjz&rsv_dl=gh_pc_zhidao" \t "/home/foreverhappy/文档\\x/_blank)把学习看作试探评价过程，Agent选择一个动作用于环境，环境接受该动作后状态发生变化，同时产生一个强化信号(奖或惩)反馈给Agent，Agent根据强化信号和环境当前状态再选择下一个动作，选择的原则是使受到[正强化](https://www.baidu.com/s?wd=%E6%AD%A3%E5%BC%BA%E5%8C%96&tn=SE_PcZhidaonwhc_ngpagmjz&rsv_dl=gh_pc_zhidao" \t "/home/foreverhappy/文档\\x/_blank)(奖)的概率增大。选择的动作不仅影响立即强化值，而且影响环境下一时刻的状态及最终的强化值。

强化学习不同于连接主义学习中的[监督学习](https://www.baidu.com/s?wd=%E7%9B%91%E7%9D%A3%E5%AD%A6%E4%B9%A0&tn=SE_PcZhidaonwhc_ngpagmjz&rsv_dl=gh_pc_zhidao" \t "/home/foreverhappy/文档\\x/_blank)，主要表现在教师信号上，强化学习中由环境提供的强化信号是Agent对所产生动作的好坏作一种评价(通常为标量信号)，而不是告诉Agent如何去产生正确的动作。由于外部环境提供了很少的信息，Agent必须靠自身的经历进行学习。通过这种方式，Agent在行动一一评价的环境中获得知识，改进行动方案以[适应环境](https://www.baidu.com/s?wd=%E9%80%82%E5%BA%94%E7%8E%AF%E5%A2%83&tn=SE_PcZhidaonwhc_ngpagmjz&rsv_dl=gh_pc_zhidao" \t "/home/foreverhappy/文档\\x/_blank)。

强化学习系统学习的目标是动态地调整参数，以达到强化信号最大。若已知r/A梯度信息，则可直接可以使用[监督学习](https://www.baidu.com/s?wd=%E7%9B%91%E7%9D%A3%E5%AD%A6%E4%B9%A0&tn=SE_PcZhidaonwhc_ngpagmjz&rsv_dl=gh_pc_zhidao" \t "/home/foreverhappy/文档\\x/_blank)算法。因为强化信号r与Agent产生的动作A没有明确的函数形式描述，所以梯度信息r/A无法得到。因此，在强化学习系统中，需要某种随机单元，使用这种随机单元，Agent在可能动作空间中进行搜索并发现正确的动作。