P22.拥有的越多，在风险中损失的也就会越多。

风险可能会带来损失，往往会多于随机中的收益。

当所处的位置较低时，向下有限，向上无限，波动所造成的平均收益为正。

奴隶们可以丢掉的，只剩下枷锁。

P23.恰当地把握一类事物，明晰其轮廓和精要，而非死缠烂打。

斤斤计较的损失，要大于收益，会丢失掉很多的机会。

对于无心学术圈的人，如果能够巧妙地把成绩控制在及格线，大学毕业后，就可以用节省出来的时间所换来的“其他技能”，做成自己想做的事情。

P25.当人口增长，超过了食物的增长，过分消耗了储量的时候，人类不可避免地陷入了饥荒、疾病、战争。

普遍性的贫穷，是人类的宿命。

P27.定点，类似于系统变化的宿命。

定点虽然存在，但只有在极特殊的条件下，才能达到。对映在人类社会的阶层状态：逆袭。

逆袭是存在的，但要极好的运气、相当高的智慧，才能勉强达到。即使达到了，稍微有风吹草动，也会岌岌可危，很可能失去这个位置。

稳定性：系统处在定点周边的状态，是否能够容错。

Balance points’ Taylor Expansion: (1) Low: first order. (2) High: Jacobian Matrix Eigenvalue

P28.处于Basin of Attraction，必归于定点。

不稳定，就是脆弱。任何的风吹草动，都能够结束表面的美丽。

全局稳定性：无论初始条件为何，系统都将趋近于这样的定点，是高度可预测系统。

P29.世界可以看作是：由不同频率振动，组成的交响曲。

为什么振动可以广泛的存在？定点是广泛存在的，振动时围绕着一个确定的状态，上下波动。

振动，西西弗斯，加缪。

Doincare-Bendixon Theorem：

Conditions：（1）2D，二维动力学系统。（2）Continue：系统连续可微。（3）Confined：动力学流在一个区域内封闭。（4）No fix point：在此区域内，定点不可达到。

Consequent：该区域内的动力学流，将收敛于一条闭合轨道->⚪

相平面的闭合轨道=周期性运动=振动

有限二维系统的运动：不存在其他情况。只有：（1）平衡态，归于定点。（2）周期性运动。

将各种不同形式的系统，都归于空间里的topo。

P30.非随机的2DS具有绝对的可预知性，2DS没有混沌。

发现振动->2个关键的动力学变量->观察是否存在平衡态。

如果没有，往往预示着：存在一个无定点的闭合区域。

P31.物理的核心：Reductionism简化。

P32.黑格尔辩证法：矛盾导致变化。

当系统有2个互相制约的变量时，就会引发振动。

Poincare-Bendixon Theory：辩证法的数学精确表达。

Lotka-Volterra Function

自然系统：初始数量，只要不是一开始就平衡，都将形成一个振动的变化。

经济学中，生产和需求作为一对相互制约的量，互相追赶却始终无法相等，于是就会陷入永不停歇的振动状态，形成经济周期。

经济运行，不可避免的，在一定时间后，走向低谷。

经济系统不是二维的，过于简化的理论，可以帮助人们理解经济现象。但是，凭借着初等的模型去实践，就会很危险。

尽量不要抽象地讨论，要还原。任何一个问题，在生活中，都是具体的。

今天许多企业资不抵债，就是因为炒概念、吹泡泡容易。

精细生产：为产品定一大堆精细而具体的指标，都达到了，索命产品质量达到了预期。

谷底变山的过程：Bifurcation Point产生的过程。

处在Bifurcation Point上的小球，具有自由意志，非常敏感。一个随机的微小扰动，都可以被放大，向左或向右。

Bifurcation结束，小球落入新的平衡点，此时，已经被一个极大的负反馈束缚住了。除非有强大的能量，否则就会一直保持这个状态。

Bifurcation Point：决定性的瞬间。此时此刻，系统前途未卜，任何风吹草动，都可以使之走向截然不同的未来。

如同在高考考场上，蒙对或者蒙错，就可能去了截然不同的城市，遇到不一样的爱情。

P40.循规蹈矩的化学平衡，异常的化学震荡，可以完全统一在一个系统之内。本质都是：Hopf Bifurcation。

心脏的规律跳动，新陈代谢循环伴随一生，这是系统的振动解。在死亡来临的那一刻，停止振动，进入静态平衡。

This is Bifurcation Point, from live to death.

P41.混沌具有其确定性的方程，只是其复杂性使得看上去像是随机的、毫无规律的。

混沌的表现混乱，不同于量子的随机性。量子，是真的随机；混沌，是无法在有限算力的情况下，计算其确定解，但其确实是一个确定系统。

P43.混沌中，相邻的轨道，对初始值非常敏感。

Liyaponov indication：一步之遥🡪天壤之别；归于一处。

P44.混动，代表着自由。自由的市场，最后会产生垄断（定点），但垄断的格局却不会太久。

自由的市场，是一个高维的混沌系统。

混沌的不可预测：描绘初始值敏感。

混沌，可以说摧毁了：人类想要预测一切的伟大理想。

P46.机器学习是：做出一个可以“自发”产生解决方案的程序，从Data中学会Program的程序。

P49.Overfitting：平时背得滚瓜烂熟，一到考试，考得一塌糊涂。

P50.Unsupervised Learning：容易产生惊喜。

我们希望：有模式、有套路，才可以精准预测。但事实上：即便找到了模式，也会在某个点上，碰到问题的边界。

P51.ML具体问题：数据清洗🡪特征提取🡪模型选择。

最难的问题是：把现实生活中的问题，提炼成一个ML问题。因为这需要对问题本身有深刻的洞见，同时又熟练掌握机器学习算法。

Auto-do：From数据清洗to交叉验证。

难点：从一个新的领域，提取出机器所要的，抽象成ML问题，以帮助人类解决最重要的问题。

如何找到人类care的问题，是机器办不到的。

用ML解决问题，需要具备三个约束：（1）系统中，可能存在模式。（2）这种模式，不是一般解析手段可以猜到的。（3）数据是可以获取的。

元素分析🡪系统预测；模式识别🡪系统聚类

P51.ML得到的：Max a Posterior；概率密度与实际情况的最佳匹配。

P52.构建大框架：调参。在某型框架下，寻找已知数据的最佳匹配的假设。

利用已知模型，在新数据里做预测、做决策。

适应问题范畴不同，模型的容纳力也不同：描述能力不同。

引入先验，来约束模型的空间大小。这需要：在偏差和误差之间平衡。

P53.Perre Laplace: Probability theory is nothing, but common sense reduced to calculation.

外部世界与内心想法之间，存在着某种交互关系。

先验概率：在取得证据之前，指定向的。通常是根据“经验常识”计算得到的，带有一定的主观色彩。

P54.不能过早就下论断，要给各种假设留下一定的空间。

·预测方法：

1. 根据已有的经验、知识、常识等，推断出一个先验概率。
2. 在新的证据，不断涌现的情况下，不断调整先验概率。

贝叶斯分析（Bayesian analysis）：累计证据🡪得到事件发生的概率

Bayesian analysis：先验、后验、条件。

P55.特例与一般常识：

1. 如果太注重特例，就很可能把噪声，当作有用的信号。
2. 如果恪守先验概率，就很可能会无视变化，墨守成规。

贝叶斯决策（Bayesian Decision）就是在不完全情报下，对部分未知的状态用主观概率估计，然后用贝叶斯公式对发生概率进行修正，最后再利用期望值和修正概率做出最优决策。

贝叶斯决策属于风险型决策，决策者虽不能控制客观因素的变化，但却掌握其变化的可能状况及各状况的分布概率，并利用期望值即未来可能出现的平均状况作为决策准则。

贝叶斯决策理论方法是统计模型决策中的一个基本方法，其基本思想是：

1、已知类条件概率密度参数表达式和先验概率。

2、利用贝叶斯公式转换成后验概率。

3、根据后验概率大小进行决策分类。

Bayesian Decision：数据、假设、目标、决策。

P56. Bayesian Decision的步骤：

1. 厘清因果链条。剥出：假设、证据。
2. 构建假设空间。
3. 开出先验概率分布。
4. 求出后验概率分布。
5. 求出条件期望，找出最大值。

Pure Bayesian Decision：一组离散的，且独立的分类问题。

P61. Decision tree：前对后，有制约、有影响。

Deepth：减少不确定的程度。

P62.可能的结果越是多，概率分布就越是均匀，信息熵越大，信息量就越没有价值。

P63.决策树：每一个分叉下，都是一个样本类，纯度高。

优点：可解释性强，可以当作：特征值的重要性排序的筛选法。

看前几层，就会对问题一目了然。

缺点：受噪声影响大。

构建集群模型：随机深林、梯度提升……串行方法等。

P67.不同的维度的数据，可以互相独立验证、互为因果，从而带来数据间的有序性。

数据在不断的变化中，变化的趋势，又受制于：数据之外的环境。环境因素，又具有自我指称性。

Data clean：是最耗时耗力的。如何避免过度补偿，如何补偿丢失等等。

P68.癌症患者：我能有几年？

终极降维，将会扭曲很多的信息。

问问题的人，不是在做科学研究，其只是想：基于短小、明确的信息，来做出准确的判断。

有用的信息，必须要在1min内说完。

对于一个随机变量，重要的不是其统计指标，而是分布。

正态分布，大体上是对称的。没有长尾效应，就不会出现过于离谱的事情。

指数分布，具有长尾，原理平均值的事件，注定会发生。（墨菲斯托定律）

Scale free：活了一千年，还能再活一千年。

P69.降维🡪给出排名🡪给出评价

然而，自然界的进化，并非如此。

不同的基因，是不同的维度。每一种基因的组合，都会带来不同的适应性。进化的目的，是找出最优的基因组合，使生物的能量利用效率越来越高。

物理的约束，使得真实世界的维度有限。看似很多，实际上，很多因素是相互依赖的。

P70.减少过度降维的弊端：找出互相独立的维度，然后通过现实中的竞争，动态调整，在不同的选项之间权衡。

生存🡪繁衍

Exploitation🡪Exploration

不能用一个简单的标准来评判，因为人是复杂的。长期、单一的用同一指标，久而久之，对于思维体系中就只剩下：单一的指标了。

具体问题，可以有一个解法；同类型的问题，一般来说，不存在唯一的范式。

P71.选择的维度：（1）构建一个相互作用的网络。（2）找出中心节点。

节点的互动，所包含的隐藏信息，可以在网络中被表现出来。从而使降维所依据的信息，不只是变量之间的相关性。

需要：来源的多样性，连续完整，稳定可靠。

P77.ReLU，Max（0，z），非线性函数中，最接近线性的，保真好。

P89.CNN：网络越深，能耗就会越大。

P91.RNN：把历史的输入，合理化为此刻决策所用的特征、记忆。

P93.神经元之间，不断交流，形成循环。但是，有路径依赖。

有一个未知的过程，有输入、有输出，通过RNN可以自己学会规律。

RNN有点Turing完备。

P94.作为一个非线性的系统中的RNN，可以继承系统内的非线性特性。

一旦RNN学习到了真实数据背后的动力系统性质，页就掌握了过程中的复杂的路径依赖，从而对过去、现在来构建模型。

P95.HMM能干的，RNN也能，不过RNN的维度会更高。

RNN用网络表达了HMM的跃迁矩阵。

HMM通过EM来自最大后验概率的算法，得出隐变量和跃迁矩阵的可能值。

RNN通过一般梯度回传算法训练。

P101.经验的积累：通过常年的试错，来计算每个行为的结果。

看书；跟高手交流，直接学习他们的经验。

Reinforcement Learning：Agent根据环境的Reward，调整action的反馈系统。

目标：Reward最大化。

难点：Agent的action改变environment；environment的改变，又返回来影响action、策略等等。

P102.人为优化，使得P(s|a)所得到的R(s)最大。

MCTS：选择、扩张、模拟估计、回传。

P103.通过看棋谱，取得一个在任意局面下，策略和最终胜率的对应关系。

直觉是需要学习的，不是天赋异禀的。当然，天赋异禀的也有，只不过很少罢了。

P108.机器学习的难点：

1. 奖励时间的不确定。每一次决策，都无法获得及时的反馈。今天有奖励的事情，明天就是悲剧。比起监督学习的实时都能得到答案，这个信息太弱了。
2. 探索与收益的平衡，无法精准把握。有的人一辈子抱残守缺，有的人一辈子都在探索，但终究都是沦为庸庸碌碌。

P112.可以设计一个，类似于地球的物理环境，让配备有深度强化学习系统的虚拟生物，任意探索。这样，可以模拟生命，然后创造上帝。

P118.大脑深层神经细胞的搜索，很难用电路去模拟，其中有各种量子事件发生，是真的随机。

以概念为目标的搜索，越是抽象，越是高效。海马体🡪空间导航。

P120.为什么古希腊的毕达哥拉斯，觉得数学的世界才是真的世界？到底是抽象的，还是具体的？是人类的幻想么？

P122.扩散的发生：能量尽可能的小，熵尽可能的大。

一切生物行动的宗旨：最大化生存可能性。Maximize existence.

外部世界，充满了危险性。如果把生物的所有行为的可能性，做成一个状态空间。那么，只有极少数可以生存。所以，生物就只能绑定在，这极少数的状态里。（生物，普遍都是逆热力学第二定律的。）

P123.作为具有认知能力的动物，最能够保证生存的“最有利的状态”就是：不断学习。

P125.一般来说，知道发现已经证实的，与自己的想法不符合，就会不断探索、不断改变环境，一直到：出现的结果，都符合自己的预期。

学习的目的，就是：让内在状态的模型，可以更加精准预测外部的环境。

模型的预测都不准确，行为决策，自然无法达到预期结果。

认知过程：对外部环境的一种编码，建立内部与外部的“一一对应”的映射关系。

P127.机器学习的一大类：模仿自然。

在大多数时候，人类如果在一些方面做得很好，那往往是仿效大自然固有的机制。

P132.在进化中，逐步发展出：对世界本身的深入理解，对理解本身的无限热爱。

P136.先做个四不像，然后让其学习猫咪的行为。当真实到不可区分的时候，就可以认为“这个四不像”就是猫。

P139.真实世界🡪表象的算法、符号世界🡪物理世界

现实🡪抽象（低维🡪高维）

神经网络建立了真实世界的模型：能够在高维空间中，找到相应的低维嵌入。

P140.资源有限🡪稀缺（物质、时间）🡪人性贪婪，所以只有具有“竞争动力”的个体，才能会存在，并把其基因传递下去。

P141.通过经济学杠杆，实现资源的优化配置：用有限的资源，办到最多的事情，并把事情办漂亮。

大脑：用有限的能量，产生更多的信息、储存更多的信息。

市场，通过给买卖双方，排出“使用、占有”资源的优先顺序。让最想付出代价的人，优先获得使用权。

对于外部信息，大脑是起到了市场的作用。大脑通过情绪，给经过的众多任务（信息源）排出优先级，合理分配能量。

市场，喜欢给“有信誉”的交易，优先；大脑，喜欢给熟悉的事物，优先给予能量。

容易沉溺于舒适区，这是大自然优化配置的表现。