

ChatGPT成功的关键因素与对图情的影响

化柏林

北京大学信息管理系

报告提纲



1

从神经网络到深度学习

2

从DL到生成式预训练

3

ChatGPT成功的关键因素

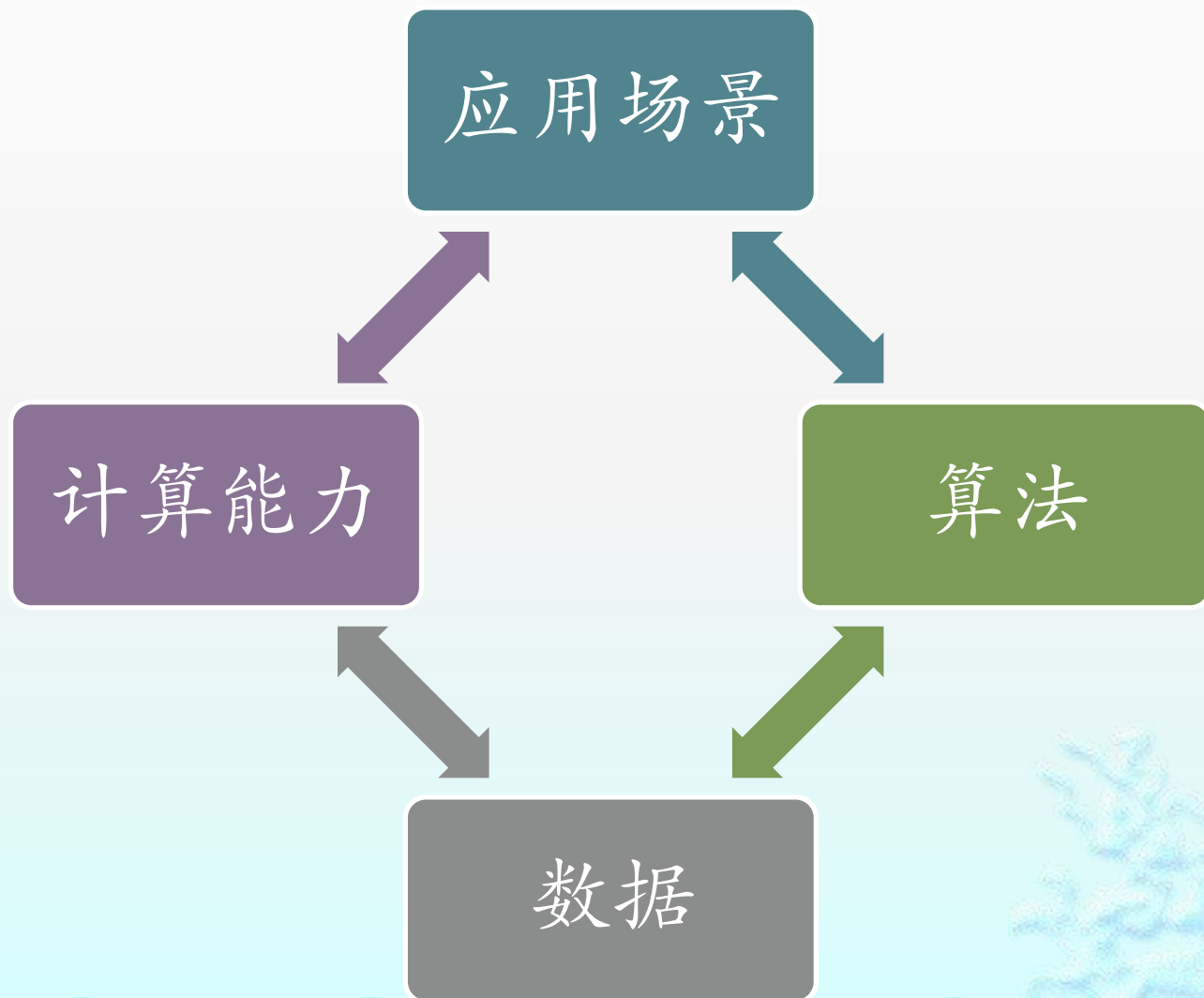
4

ChatGPT对图书情报的影响

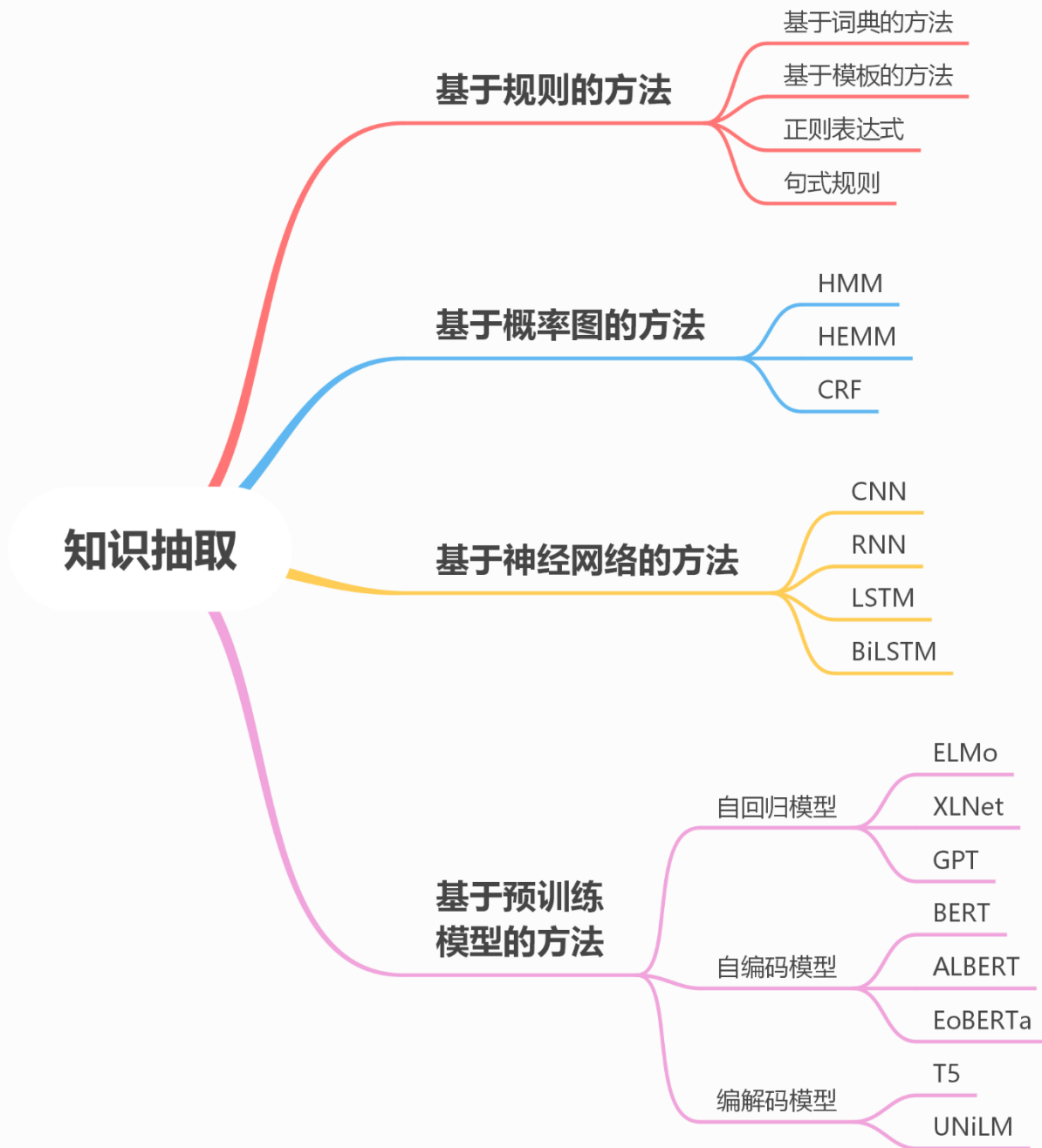
人工智能的三大学派

- ◆ 符号主义
 - ◆ 研究抽象思维、注重数学可解释性
- ◆ 连接主义
 - ◆ 研究形象思维、注重仿人脑模型、更加感性
- ◆ 行为主义
 - ◆ 研究感知思维、偏向于应用和模拟

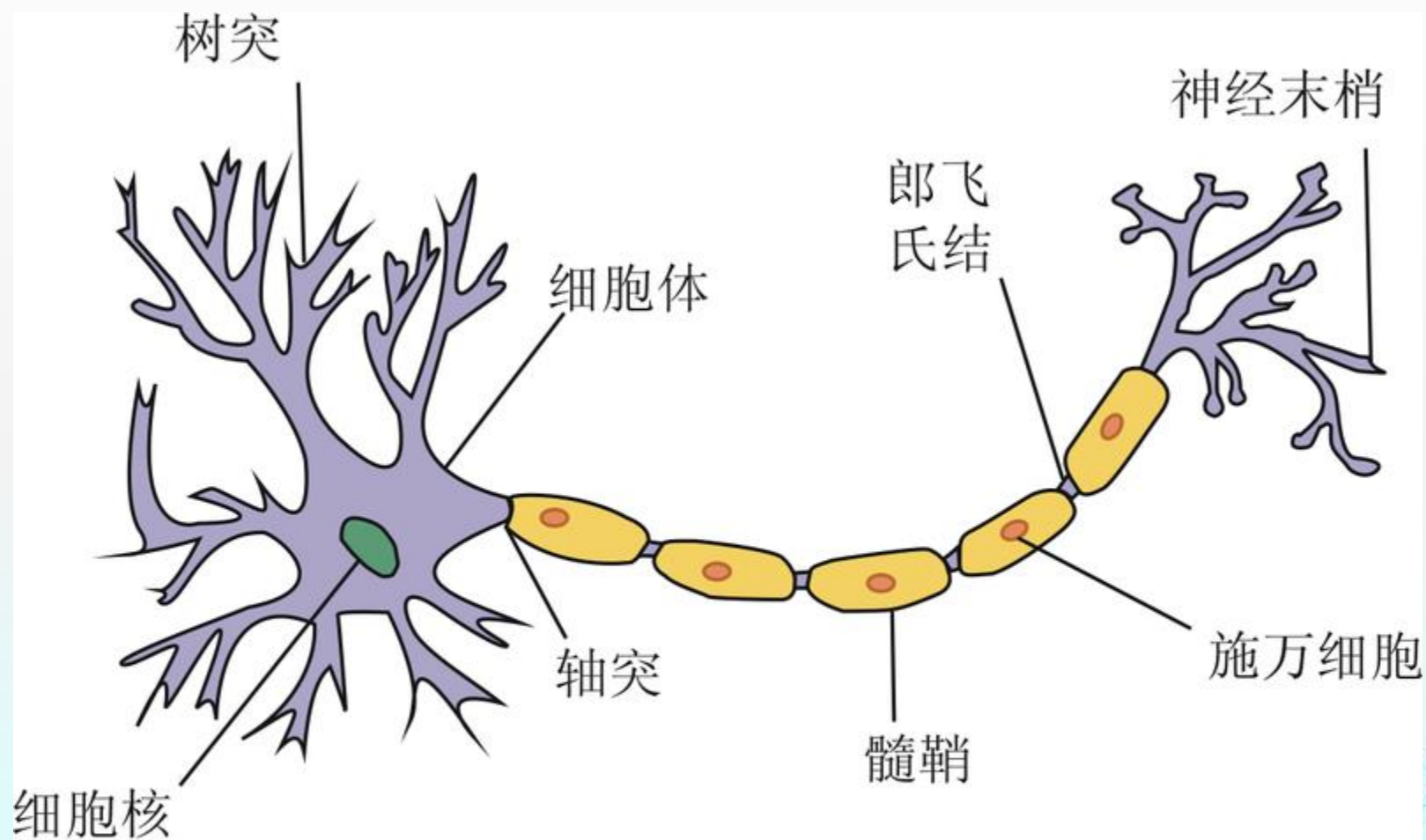
深度学习实践的四个关键要素



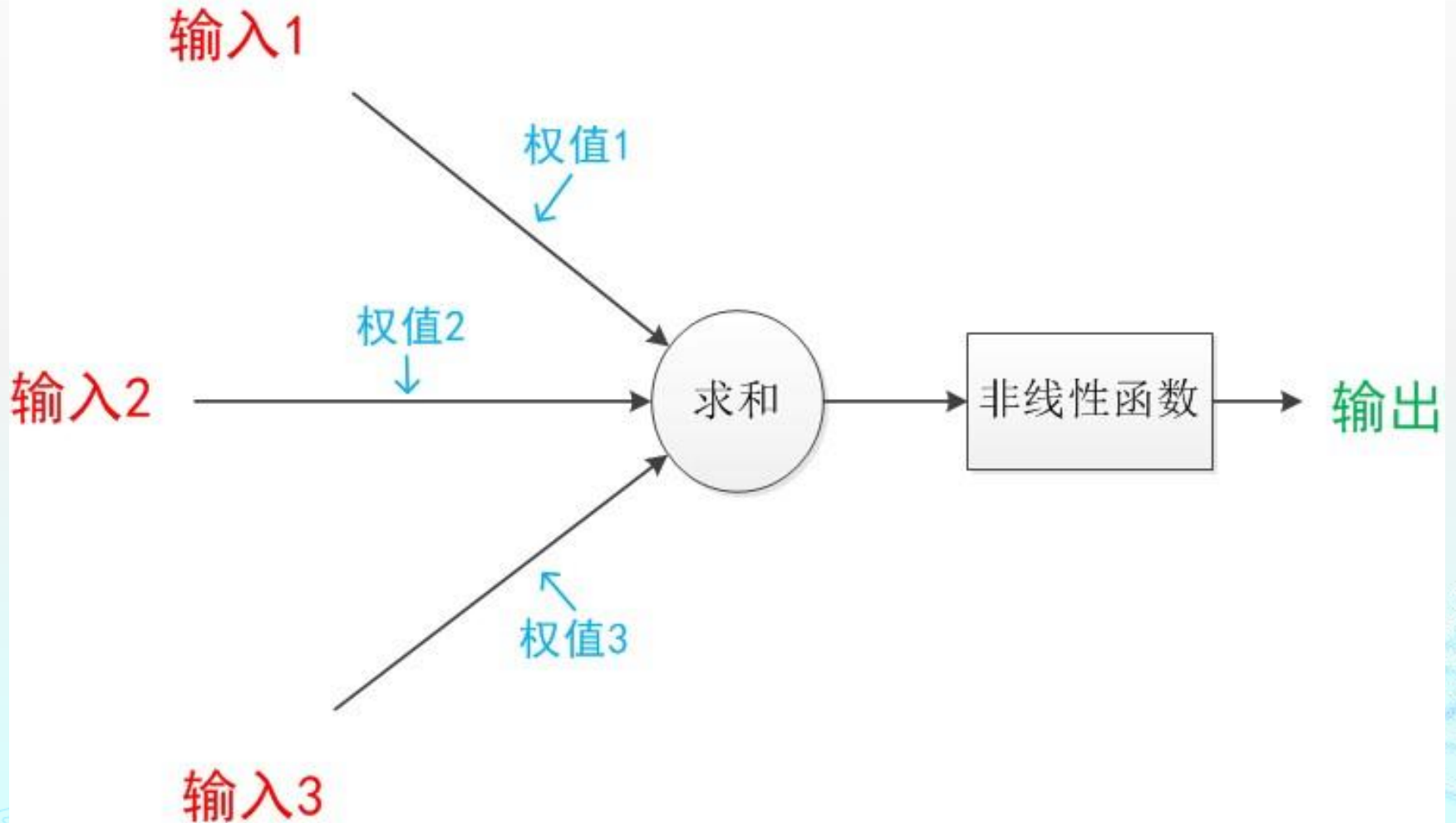
AI应用于NLP的技术路线



神经元结构

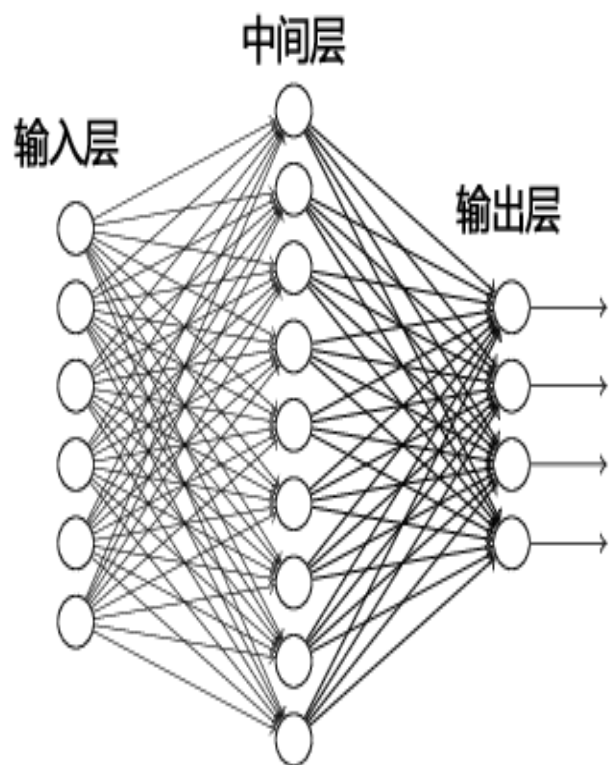


神经元模型

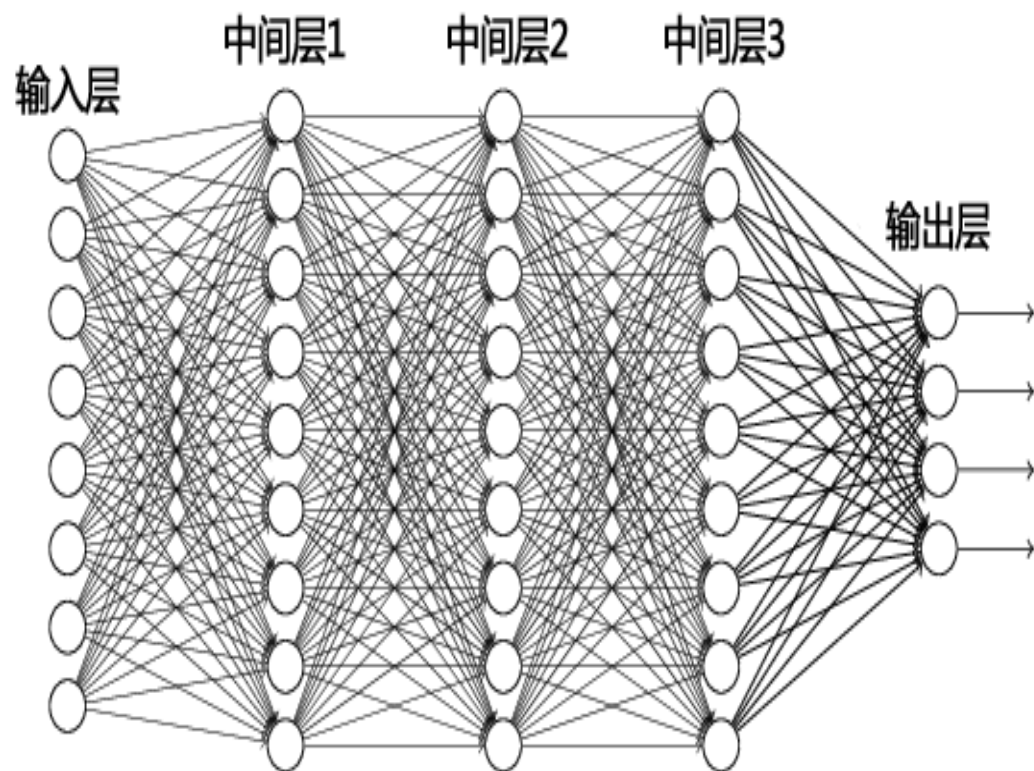


深度神经网络

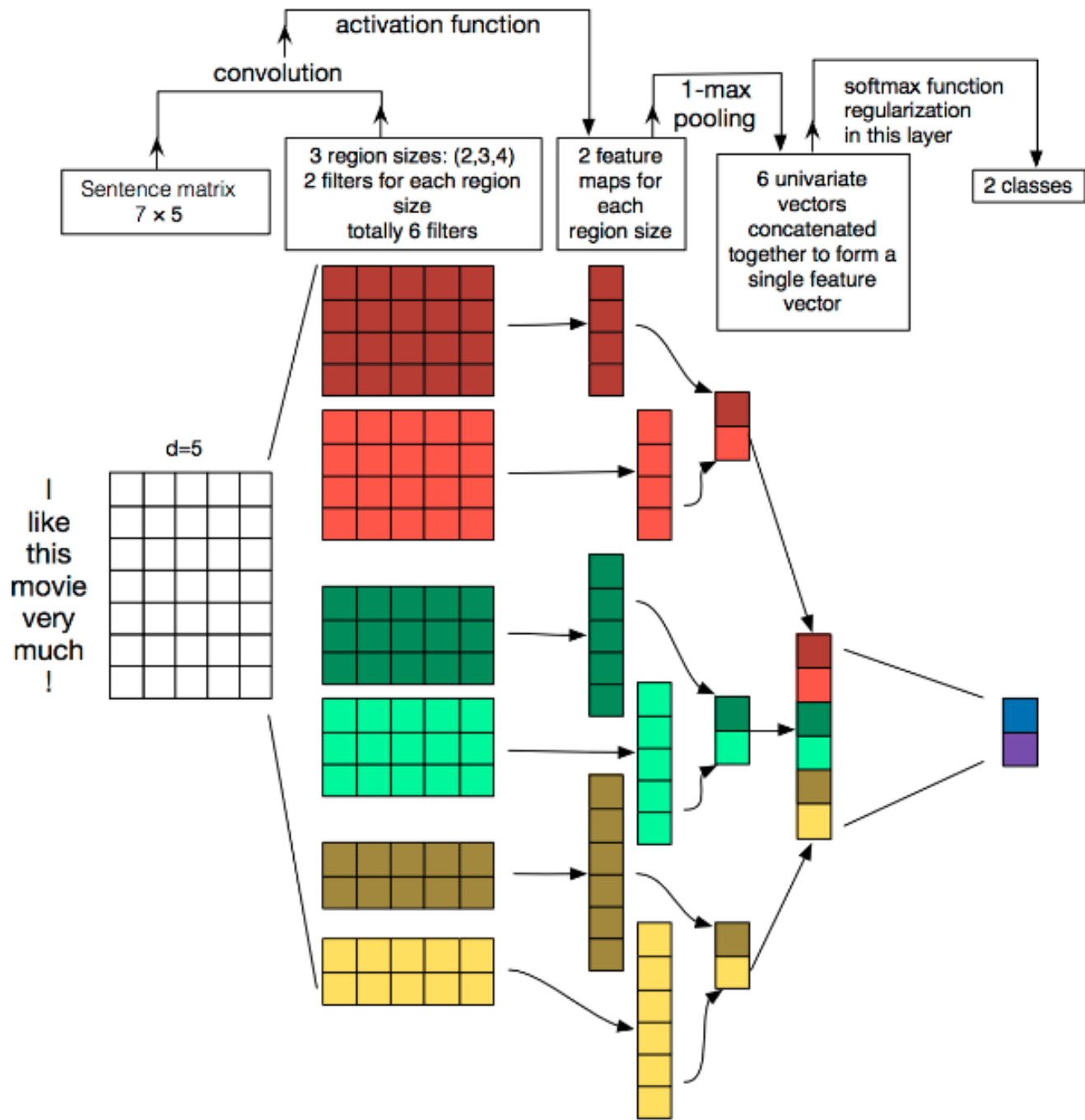
“非深度”神经网络



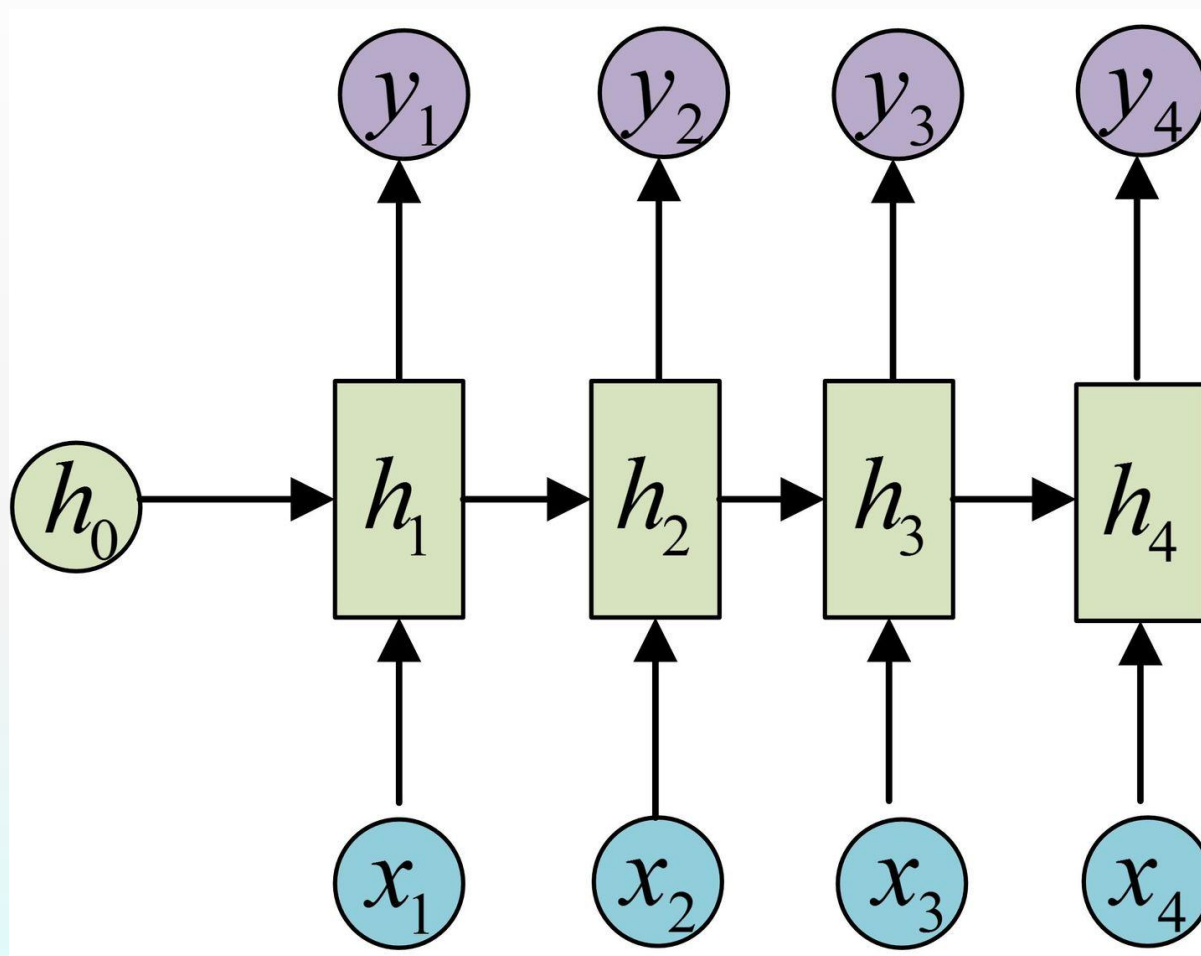
深度神经网络



卷积神经网络 CNN

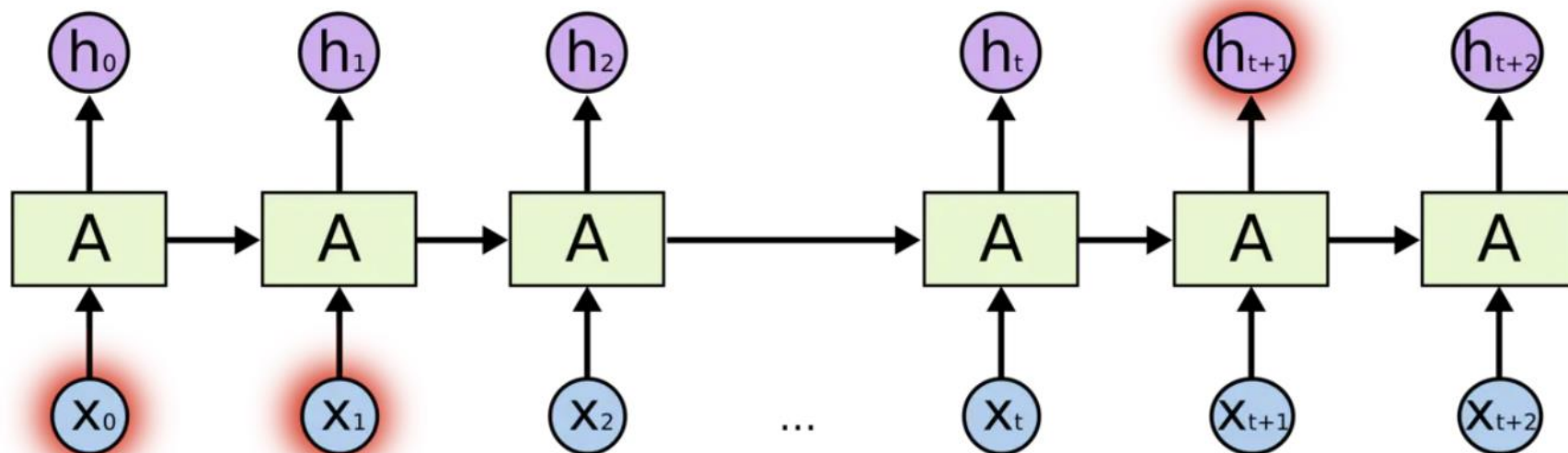


循环神经网络RNN



- ◇ 输入是 x_1, x_2, \dots, x_n ，输出为 y_1, y_2, \dots, y_n ，也就是说，输入和输出序列必须要是等长的

循环神经网络

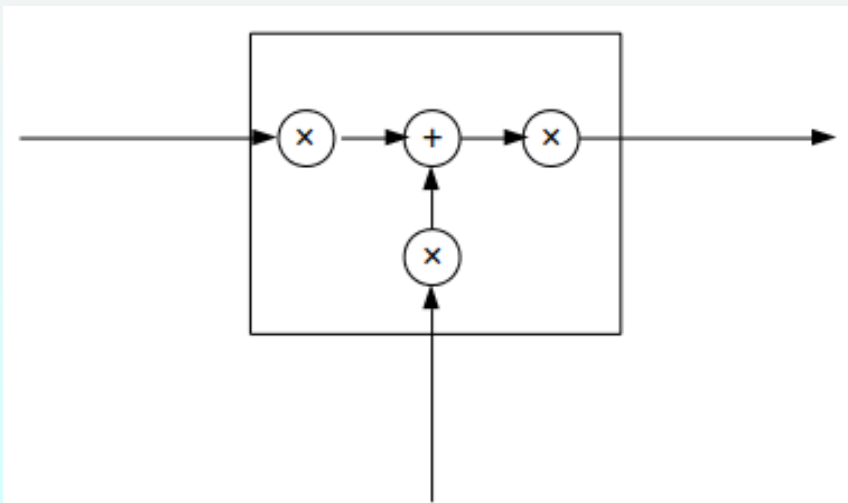


相当长的相关信息和位置间隔

I grew up in China, so I can speak fluent **Chinese**

LSTM

- ◆ 三个×分别代表的就是forget gate, input gate, output gate,
- ◆ 用forget gate, 控制hidden state信息保留多少,
- ◆ 用三个函数来控制流入流出, 输入门控制当前信息输入多少
- ◆ 门(代表sigmoid函数)的值是介于0到1之间的, 趋近于0时表示流入不能通过gate, 趋近于1时表示流入可以通过gate。



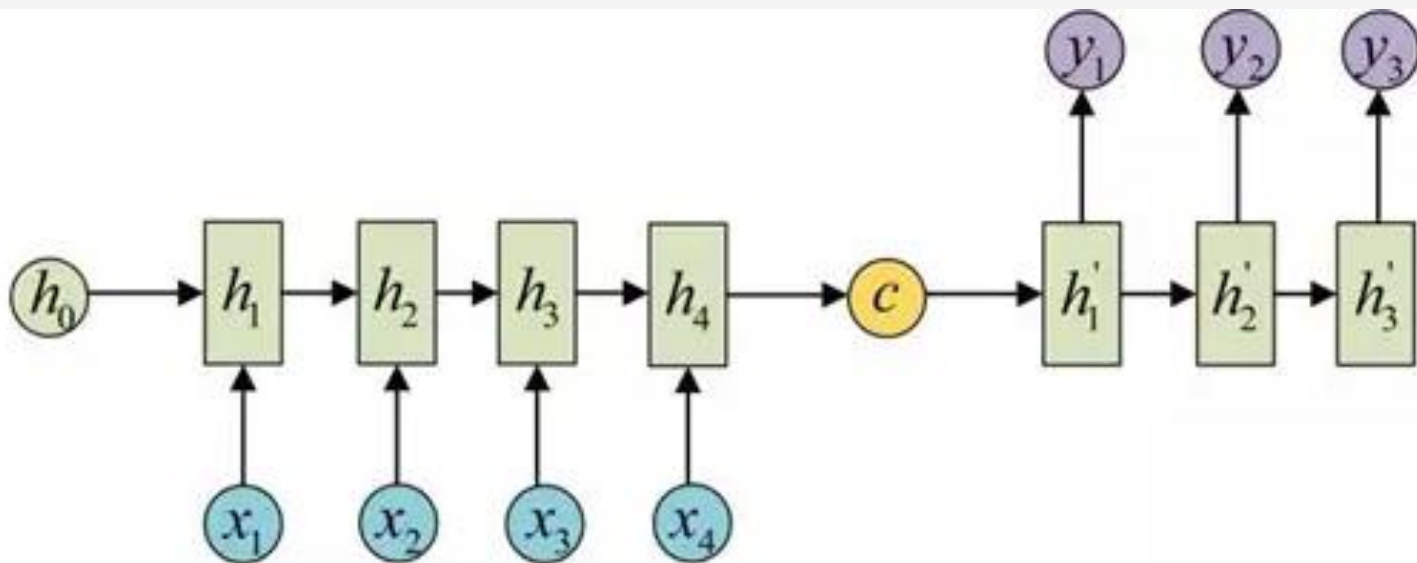
$$I_t = \sigma(X_t W_{xi} + H_{t-1} W_{hi} + b_i)$$

$$F_t = \sigma(X_t W_{xf} + H_{t-1} W_{hf} + b_f)$$

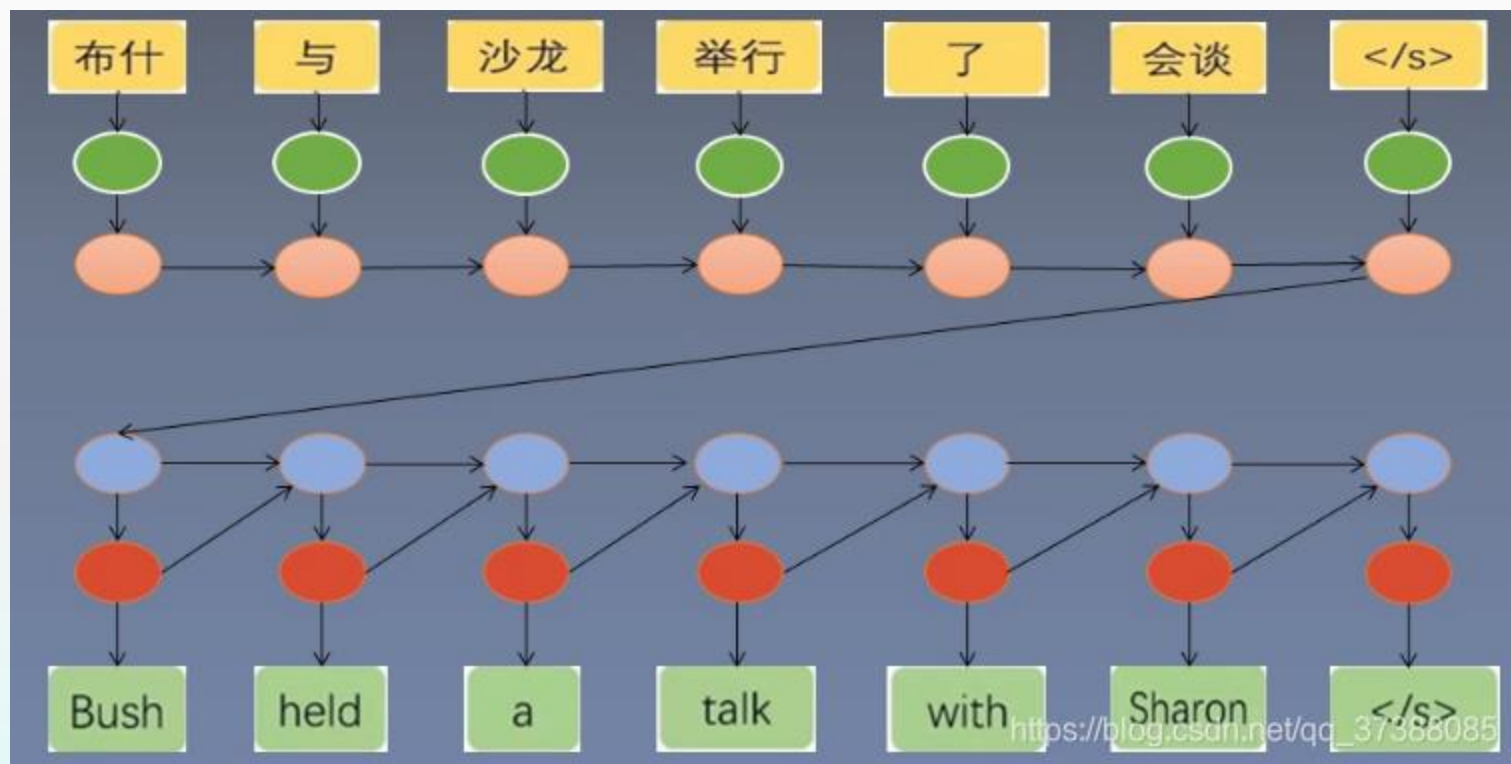
$$O_t = \sigma(X_t W_{xo} + H_{t-1} W_{ho} + b_o)$$

seq2seq

- ◆ seq2seq属于encoder-decoder结构的一种
- ◆ 利用两个RNN，分别作为encoder和decoder
 - ◆ encoder负责将输入序列压缩成指定长度的向量，该向量可以看成是这个序列的语义，此过程称为编码
 - ◆ decoder负责根据语义向量生成指定的序列，即解码



词不对齐、句序自调



报告提纲

1 从神经网络到深度学习

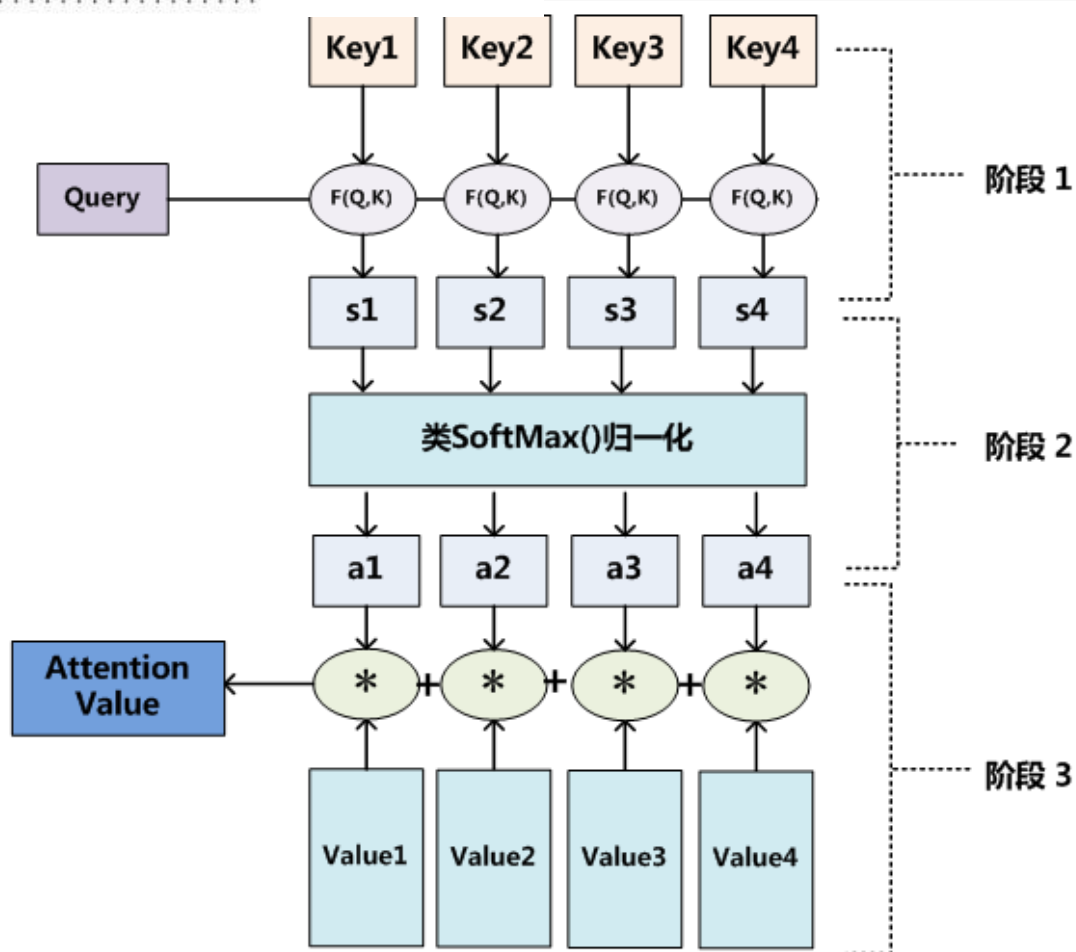
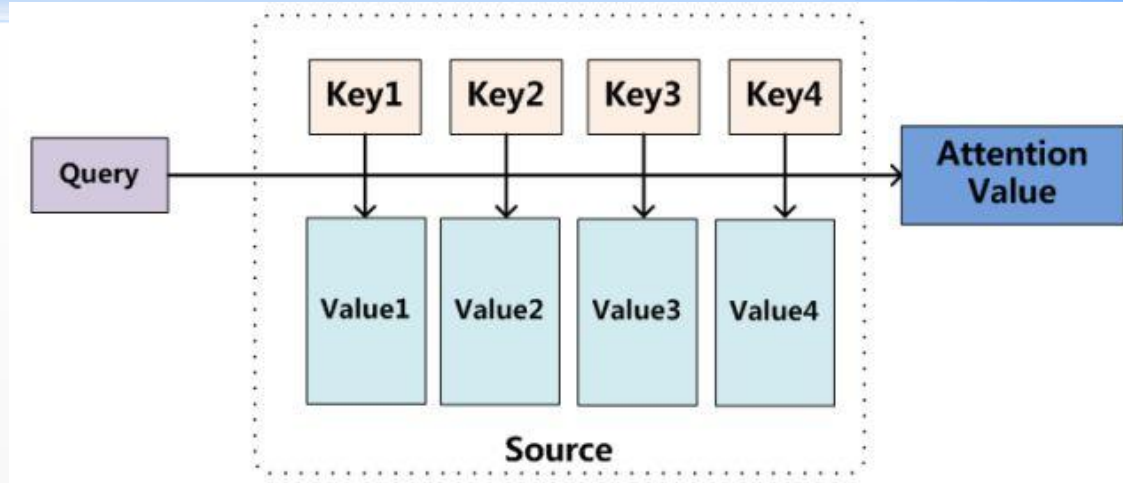


2 从DL到生成式预训练

3 ChatGPT成功的关键因素

4 ChatGPT对情报的影响

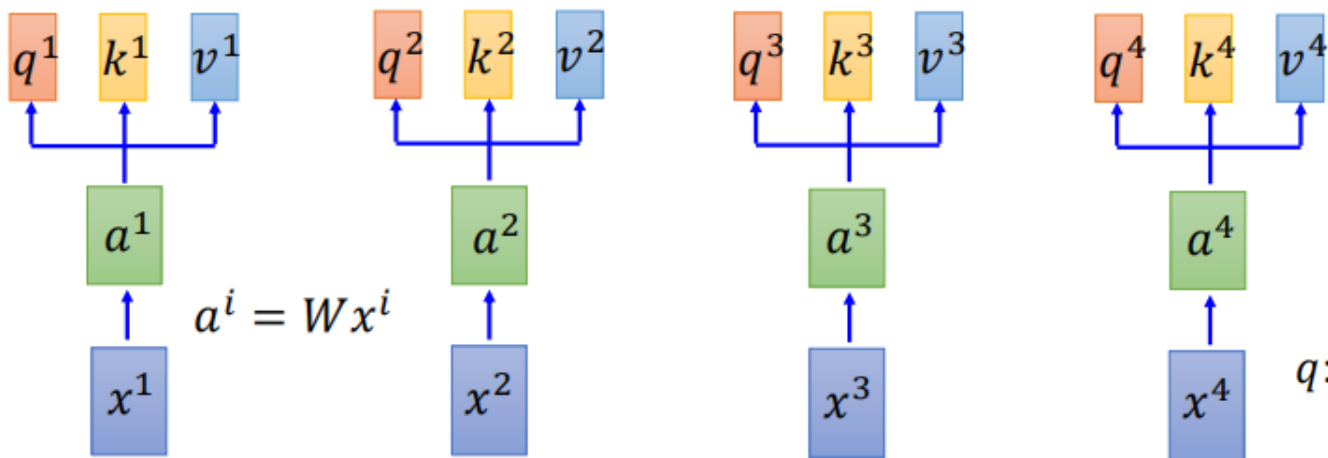
注意力机制



注意力机制

- ◆ self-attention的本质就是自己注意自己，Q、K、V来源是一样的。
- ◆ 对序列本身做 Attention，来获得序列内部的联系

$$\text{Attention value} = \text{Attention}(W_q X, W_k X, W_v X)$$



q : query (to match others)

$$q^i = W^q a^i$$

k : key (to be matched)

$$k^i = W^k a^i$$

v : information to be extracted

$$v^i = W^v a^i$$

BERT的掩码

训练数据中，15%的词被mask掉
被[MASK]替换的概率是80%，
不变的概率是10%，
随机替换的概率是10%

报告提纲

1 从神经网络到深度学习

2 从DL到生成式预训练

3 ChatGPT成功的关键因素

4 ChatGPT对情报的影响

chatGPT成功的因素

- ◆ 低维稠密向量表示文本
- ◆ 大规模预训练模型
- ◆ 微调+提示学习
- ◆ 高质量大规模训练数据

特征表示

字符级别的
嵌入表示

One-Hot编
码

词向量特征
表示

连续词袋特
征表示

生成模型视角

Topics

gene 0.04
dna 0.02
genetic 0.01
...

life 0.02
evolve 0.01
organism 0.01
...

brain 0.04
neuron 0.02
nerve 0.01
...

data 0.02
number 0.02
computer 0.01
...

Documents

Seeking Life's Bare (Genetic) Necessities

COLD SPRING HARBOR, NEW YORK— How many genes does an organism need to survive? Last week at the genome meeting here,* two genome researchers with radically different approaches presented complementary views of the basic genes needed for life. One research team, using computer analyses to compare known genomes, concluded that today's organisms can be sustained with just 250 genes, and that the earliest life forms required a mere 128 genes. The other researcher mapped genes in a simple parasite and estimated that for this organism, 800 genes are plenty to do the job—but that anything short of 100 wouldn't be enough. Although the numbers don't match precisely, those predictions

"are not all that far apart," especially in comparison to the 75,000 genes in the human genome, notes Siv Andersson, a biologist at Uppsala University in Sweden, who arrived at the 800 number. But coming up with a consensus answer may be more than just a genetic numbers game, particularly as more and more genomes are completely mapped and sequenced. "It may be a way of organizing any newly sequenced genome," explains Arcady Mushegian, a computational molecular biologist at the National Center for Biotechnology Information (NCBI) in Bethesda, Maryland. Comparing an

Stripping down. Computer analysis yields an estimate of the minimum modern and ancient genomes.

* Genome Mapping and Sequencing, Cold Spring Harbor, New York, May 8 to 12.

SCIENCE • VOL. 272 • 24 MAY 1996

ADAPTED FROM NCBI

Topic proportions and assignments

词袋的假设

1. 每个文档是由一个独特的主题分布随机生成的
2. 每个词是从这些主题中抽取的，主题包含所有词

词嵌入

1-of-N Encoding

apple = [1 0 0 0 0]

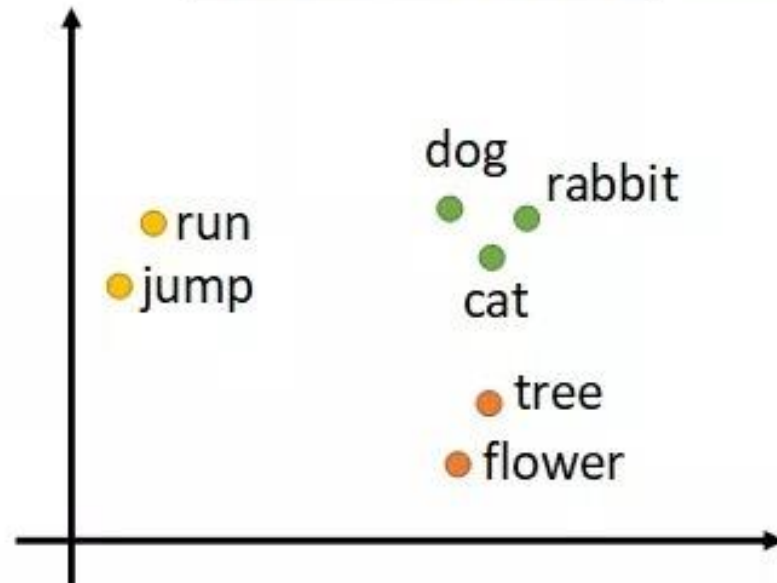
bag = [0 1 0 0 0]

cat = [0 0 1 0 0]

dog = [0 0 0 1 0]

elephant = [0 0 0 0 1]

Word Embedding



Word Class



king - man + woman \approx queen

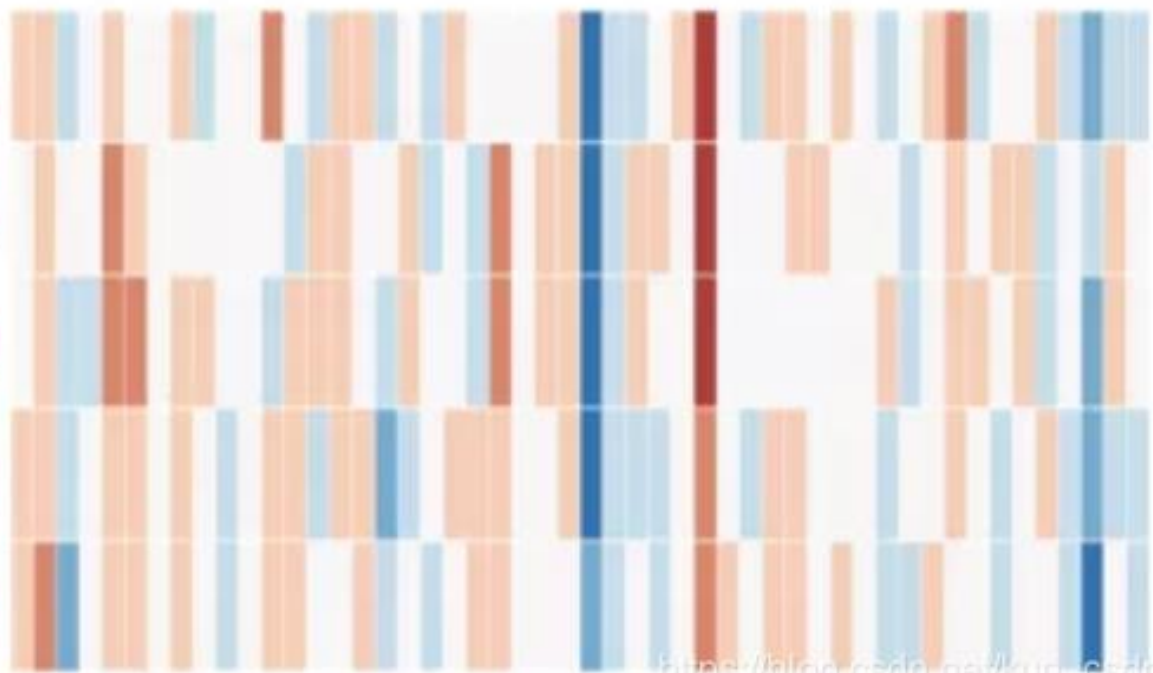
king

man

woman

king-man+woman

queen



GPT系列

- ◆ 基于文本预训练的GPT-1、GPT-2、GPT-3三代模型都采用以Transformer为核心结构的模型，不同的模型层数与词向量长度等超参。

模型	发布时间	层数	Head数	词向量长度	参数量	预训练数据量
GPT-1	2018年6月	12	12	768	1.17亿	约5GB
GPT-2	2019年2月	48		1600	15亿	约40GB
GPT-3	2020年5月	96	96	12888	1750亿	约45TB

GPT-1

- ◆ 基于Transformer Decoder预训练+FineTune
微调
 - ◆ 预训练：利用语言模型LM进行预训练学习
 - ◆ 微调：通过微调解决下游任务

GPT-2

- ◆ 舍弃微调，直接利用zero-shot learning
- ◆ 零样本学习（zero-shot learning）：没有任何训练样本进行微调训练的情况下，让预训练模型完成特定下游任务
- ◆ 单样本学习（one-shot learning）：在一个训练样本进行微调训练的情况下，让预训练模型完成特定下游任务
- ◆ 小样本学习（few-shot learning）：在只有少量样本进行微调训练的情况下，让预训练模型完成特定任务

GPT-3

- ◆ Prompt-Tuning的动机旨在解决目前传统Fine-Tuning的两个痛点
- ◆ 降低语义差异 (Bridge the gap between Pre-Training and Fine-tuning)：预训练任务主要以MLM为主，而下游任务则重新引入新的训练参数，两个阶段的目标通常有较大差异，因此需要缩小这个差距
- ◆ 避免过拟合:由于微调阶段需要引入额外的参数以适配相应的任务需要，因此在样本有限的情况下容易发生过拟合，降低了模型的泛化能力。

GPT各阶段

预训练+微调

预训练
+ZeroShot

预训练+Prompt
Tuning

预训练
+Instruct
learning

预训练
+IL+RLHF

GPT-1

GPT-2

GPT-3

Instruct
GPT

ChatGPT

ChatGPT的核心技术

- ◆ **架构：**（1）Transformer 是底层核心架构，Transformer架构有很好的并行性、可扩展性、长距离依赖等特性。
- ◆ **推理：**（2）Codex用于代码、文本生成，可以根据用户描述生成代码，而且因其上下文窗口长达8192，无形中大大提升了整个模型推理能力。
- ◆ **反馈：**（3）人类反馈强化学习（Reinforcement Learning from Human Feedback, RHFL）模型将预训练语言模型按照人类反馈进一步微调以符合人类偏好，利用人类反馈信息直接优化模型。

大语言模型的原则

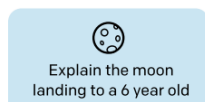
- ◆ helpful, 应该帮助用户解决任务;
- ◆ honest, 不应该捏造信息或误导用户;
- ◆ harmless, 不应该对人或环境造成物理、心理或社会伤害。

人类反馈奖赏模型

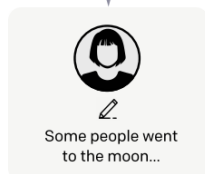
Step 1

Collect demonstration data, and train a supervised policy.

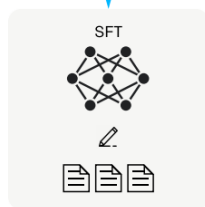
A prompt is sampled from our prompt dataset.



A labeler demonstrates the desired output behavior.



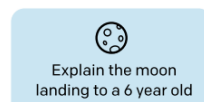
This data is used to fine-tune GPT-3 with supervised learning.



Step 2

Collect comparison data, and train a reward model.

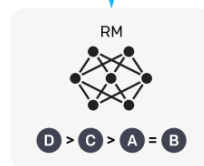
A prompt and several model outputs are sampled.



A labeler ranks the outputs from best to worst.



This data is used to train our reward model.



Step 3

Optimize a policy against the reward model using reinforcement learning.

A new prompt is sampled from the dataset.



The policy generates an output.



The reward model calculates a reward for the output.



The reward is used to update the policy using PPO.



Figure 2: A diagram illustrating the three steps of our method: (1) supervised fine-tuning (SFT), (2) reward model (RM) training, and (3) reinforcement learning via proximal policy optimization (PPO) on this reward model. Blue arrows indicate that this data is used to train one of our models. In Step 2, boxes A-D are samples from our models that get ranked by labelers. See Section 3 for more details on our method.

对齐

- ◆ 对齐技术可以针对特定应用与特定人类参考组进行对齐
 - ◆ 标注者自身的偏好
 - ◆ 标注者的代表性
 - ◆ 研究人员的偏好
 - ◆ 最终用户

InstrucGPT的后续改

- ◆ 可以尝试许多方法来进一步降低模型产生有毒、有偏见或其他有害输出的倾向
- ◆ 让模型拒绝某些用户指令
- ◆ 将 RLHF 与其他可控性方法相结合
- ◆ 其他对齐算法
- ◆ 除比较以外更有效的提供对齐信号的方式
- ◆ 减轻性能回归
- ◆ 如何设计一个透明的对齐过程，它有意义地代表受技术影响的人，并以一种在许多群体之间达成广泛共识的方式综合人们的价值观

报告提纲

1

从神经网络到深度学习

2

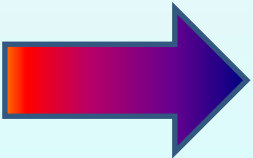
从DL到生成式预训练

3

ChatGPT成功的关键因素

4

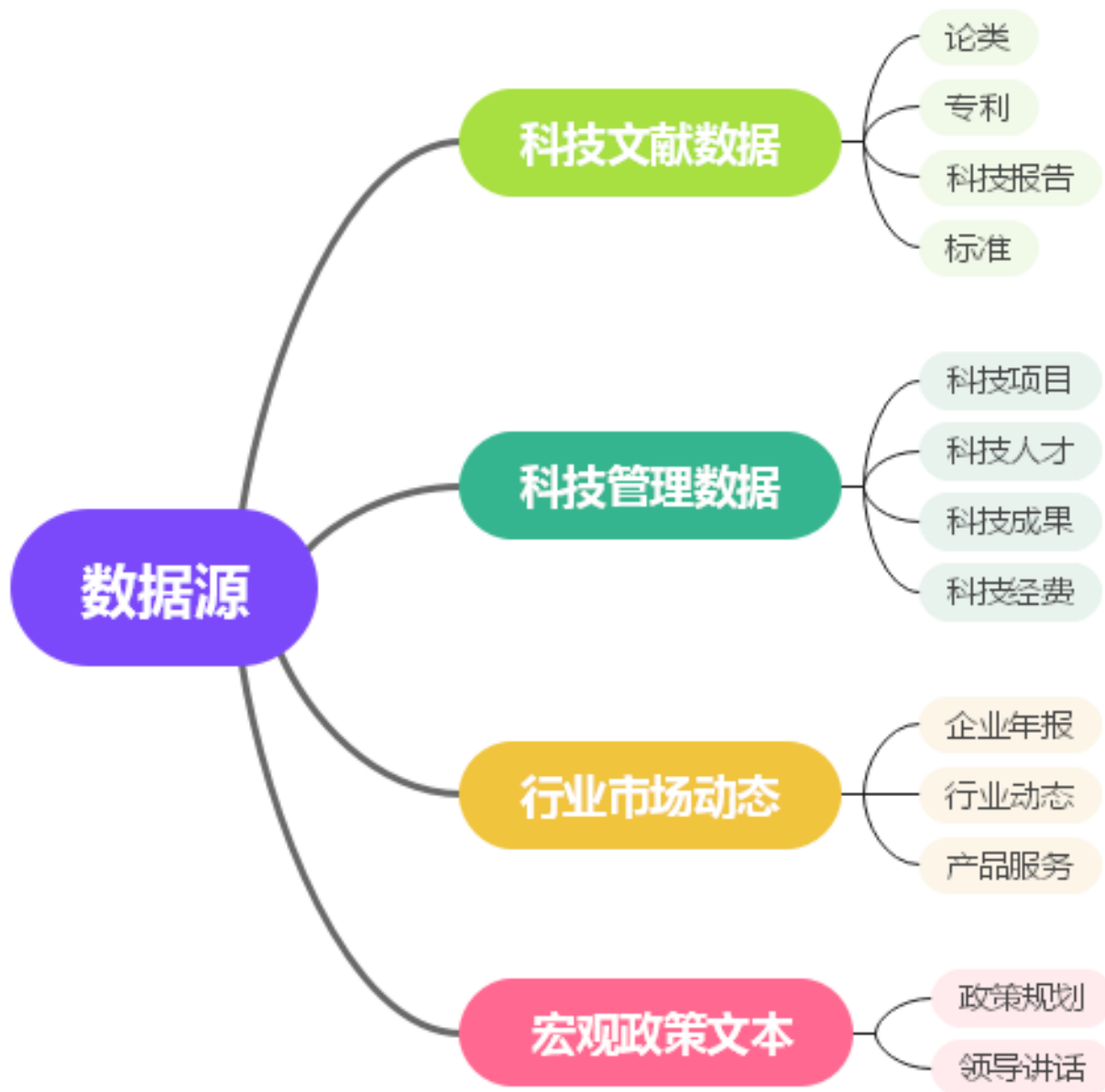
ChatGPT对情报的影响



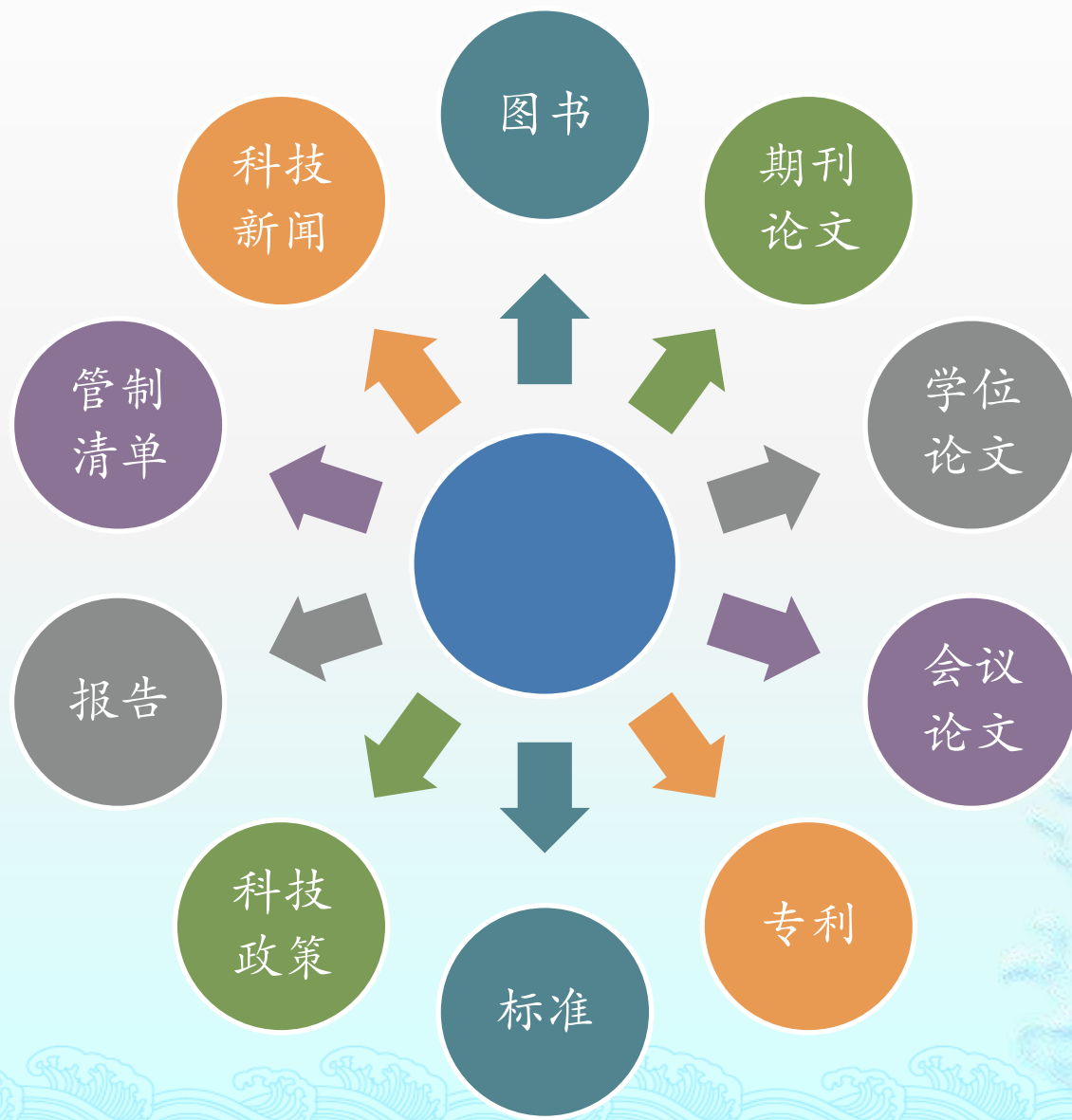
思考

- ◆ ChatGPT只是一个产品，系列的AI技术才应是关注的重点
- ◆ 强有力地支撑参考咨询、学科馆员工作
- ◆ 强有力地支撑动态类、综述类报告
- ◆ 情报从业人员整体上还难以被取代？
- ◆ 研制我们的S&T-GPT是否可行？

科技情报系统的数据源



科技文本信息的分布



谢谢大家！

- ◆ 欢迎批评指正，
- ◆ 恳请宝贵建议！