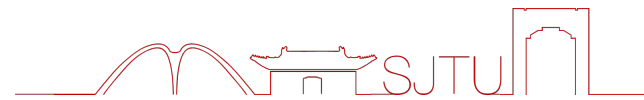




上海交通大学
SHANGHAI JIAO TONG UNIVERSITY



动手学深度学习

张家儒

上海交通大学

饮水思源 · 爱国荣校

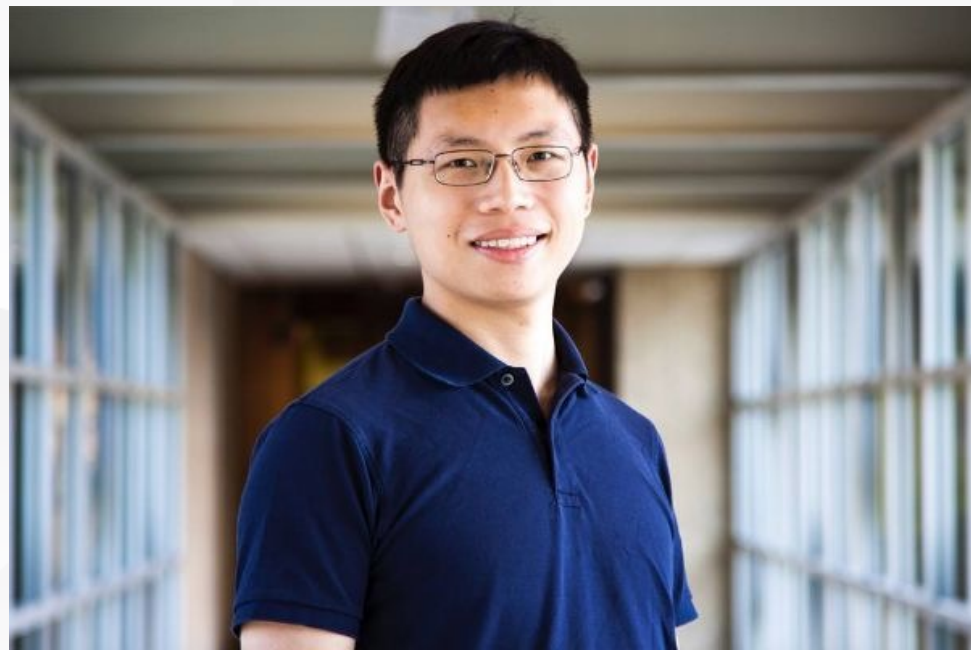
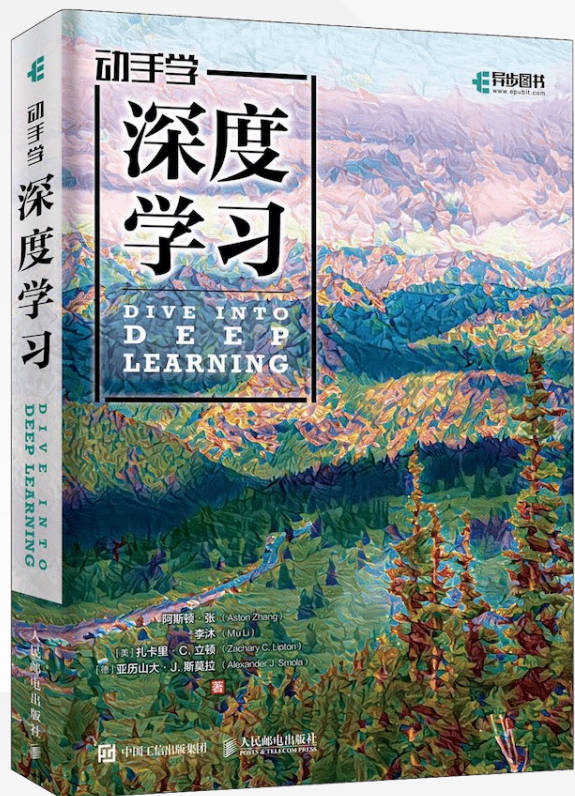


致谢



📖 在线教材: <https://zh.d2l.ai/>

📖 在线课程: <https://courses.d2l.ai/zh-v2/>





1

环境安装

2

环境安装

3

线性神经网络

4

多层感知机

5

卷积神经网络



01

环境安装



📖 教程: https://zh.d2l.ai/chapter_installation/index.html

📖 软件

- Anaconda: <https://www.anaconda.com/>
- VSCode: <https://code.visualstudio.com/>
- PyCharm: <https://www.jetbrains.com/pycharm/>

📖 Python package

- Torch
- Torchvision
- D2I



02

网络实现



① 从零实现 (代码讲解)

- 生成数据集
- 读取数据集
- 初始化模型参数
- 定义模型
- 定义损失函数
- 定义优化算法
- 训练

② 简洁实现 (代码讲解)



Softmax 回归



④ 图像分类数据集 (代码讲解)

④ 从零实现 (代码讲解)

④ 简洁实现 (代码讲解)



多层感知机



① 从零实现

② 简洁实现





卷积神经网络



① 从零实现

② 简洁实现





在 LeNet 的基础上，进行以下研究：

- 将平均汇聚层替换为最大汇聚层，会发生什么？
- 尝试构建一个基于LeNet的更复杂的网络，以提高其准确性。
 - 调整卷积窗口大小。
 - 调整输出通道的数量。
 - 调整激活函数（如ReLU）。
 - 调整卷积层的数量。
 - 调整全连接层的数量。
 - 调整学习率和其他训练细节（例如，初始化和轮数）。
- 在MNIST、FashionMNIST等数据集上尝试以上改进的网络，分析结果变化的原因，总结经验。
- 显示不同输入（例如毛衣和外套）时，LeNet第一层和第二层的激活值。尝试解释。

作业提交：实验报告+源代码

截止日期：2023-10-27 23:59





03

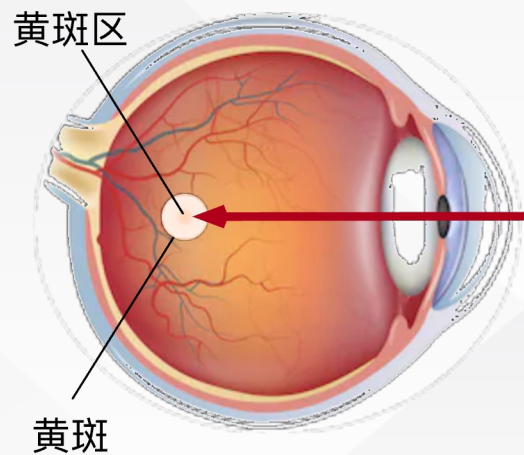
注意力机制



注意力：从生物学出发



⊙ **非自主性提示：** 基于环境中物体的突出性和易见性

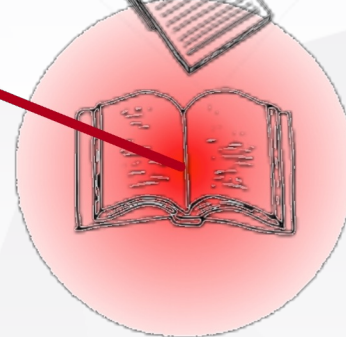
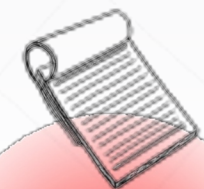
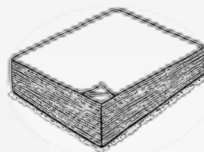
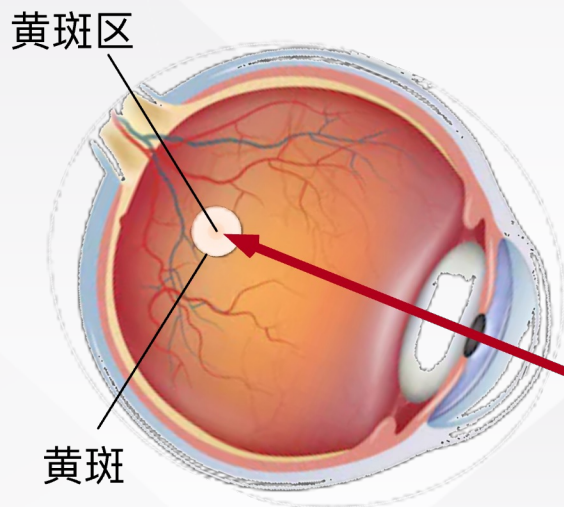




注意力：从生物学出发



⊙ 自主性提示：基于主观意愿





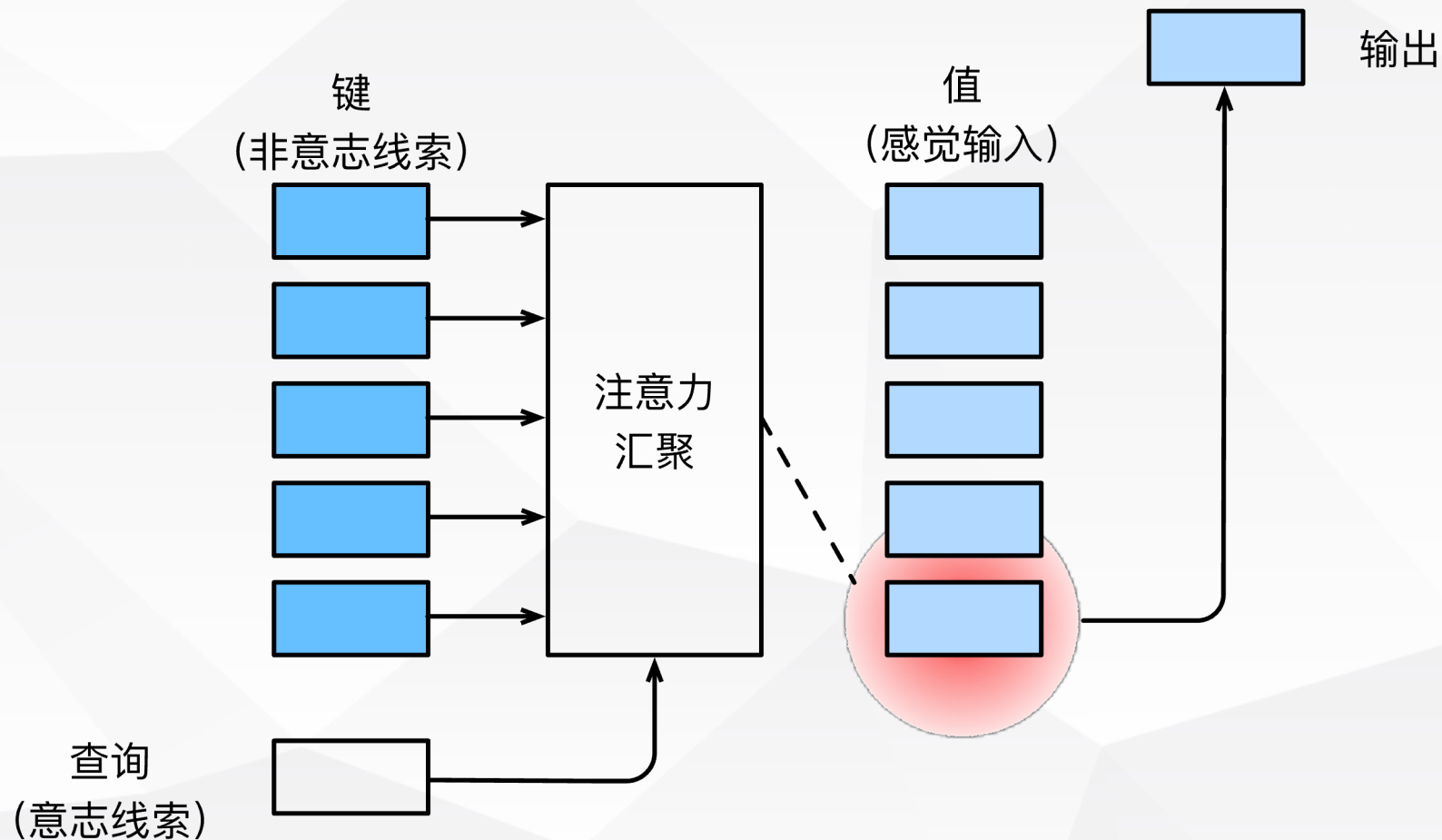
注意力：注意力机制实现方案



Query

Key

Value





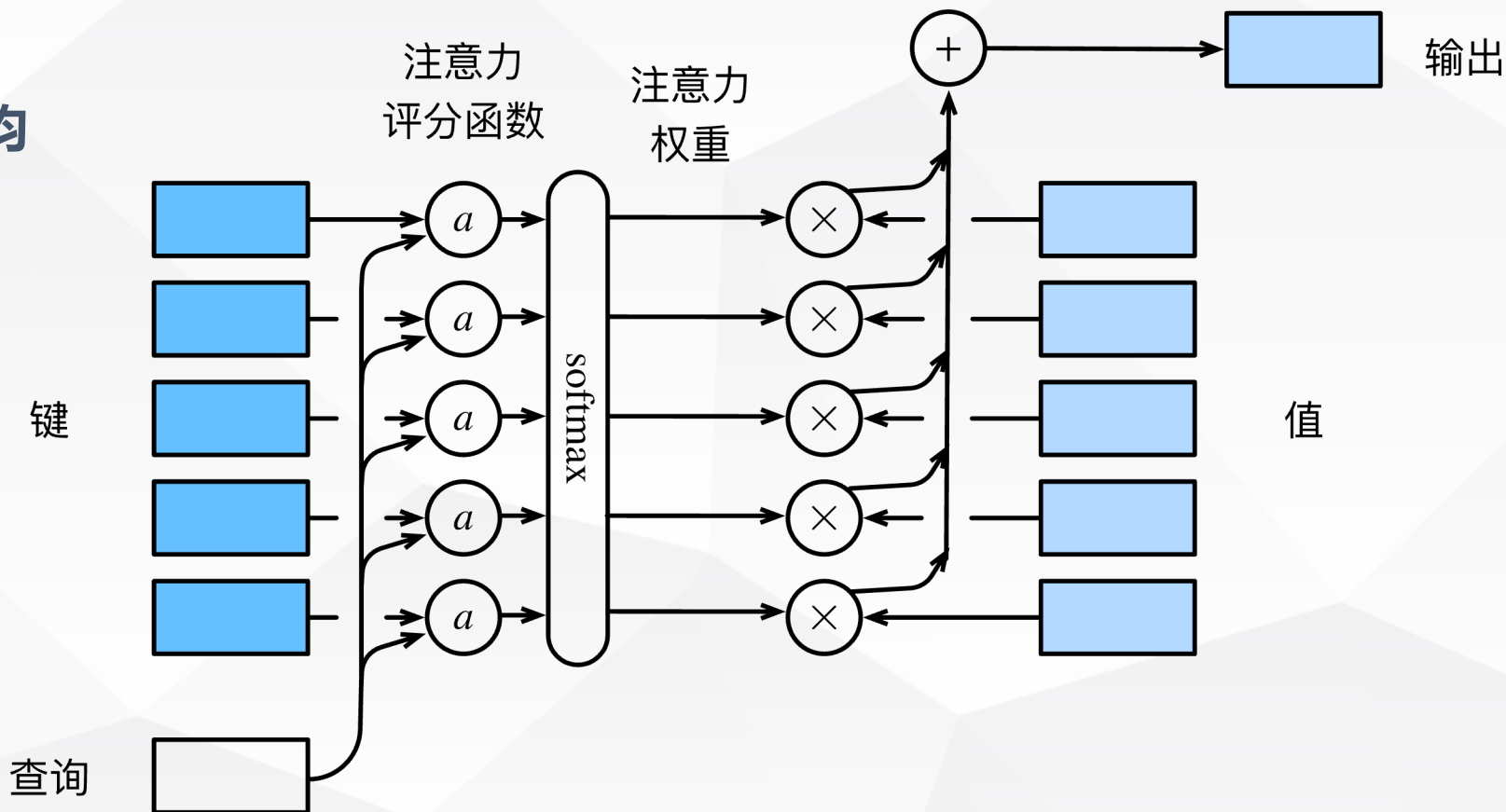
注意力汇聚：加权平均

根据Q和K计算权

使用权值对计算V的平均

常用评分函数：

$$a(\mathbf{q}, \mathbf{k}) = \mathbf{q}^\top \mathbf{k} / \sqrt{d}.$$



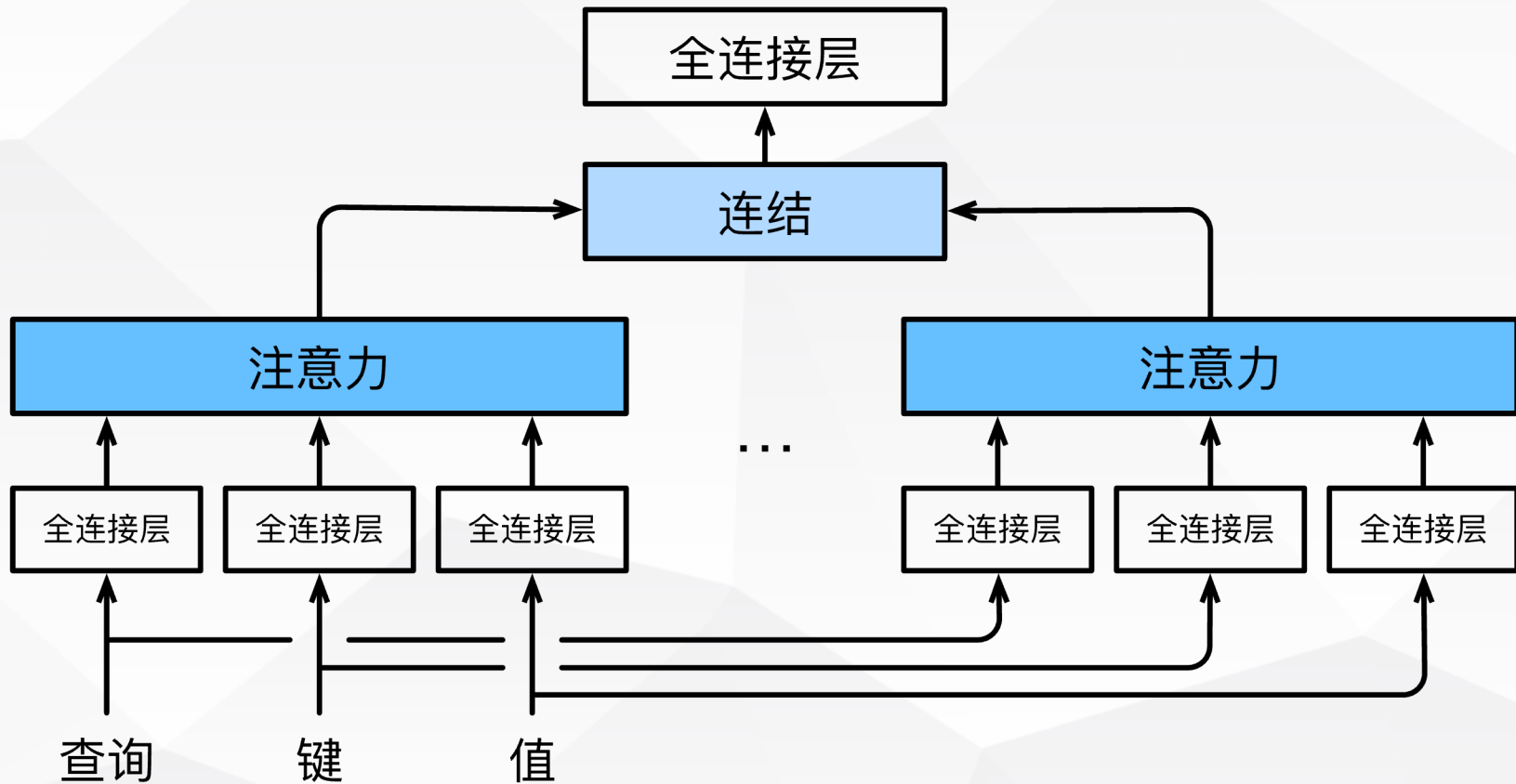


多头注意力



① 类比CNN中的多个卷积核

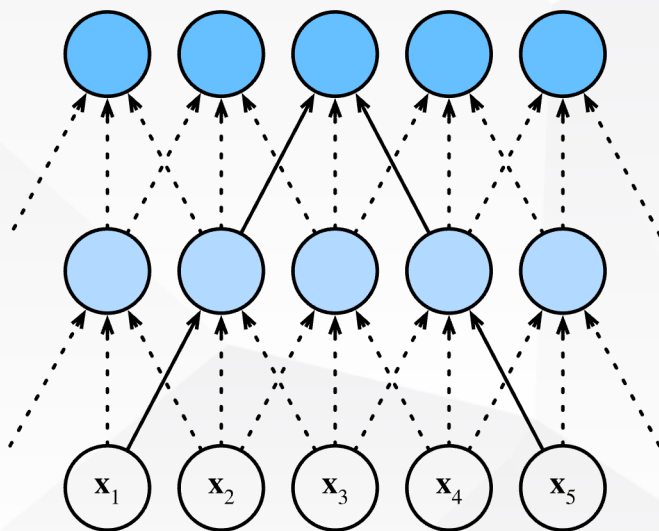
② 使用多个注意力“头”



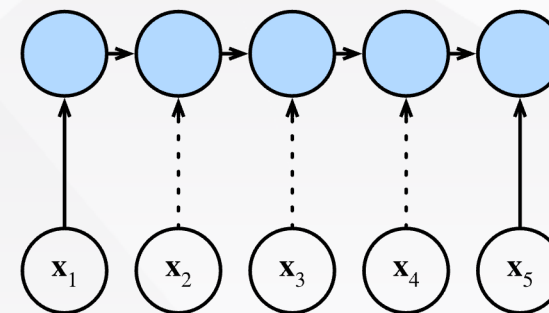


- 在多头注意力中
- 使用相同的Q, K, V输入
- 与CNN, RNN比较:
- 可并行计算
- 可捕捉远距离信息
- 计算复杂度高

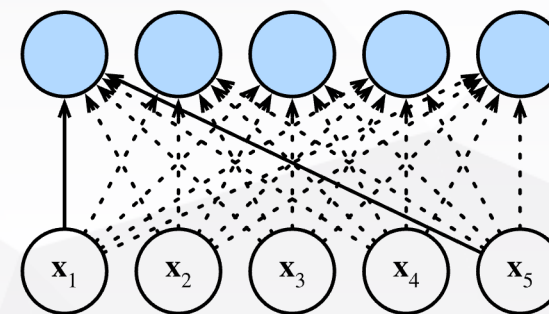
卷积神经网络



循环神经网络



自注意力





① 自注意力机制

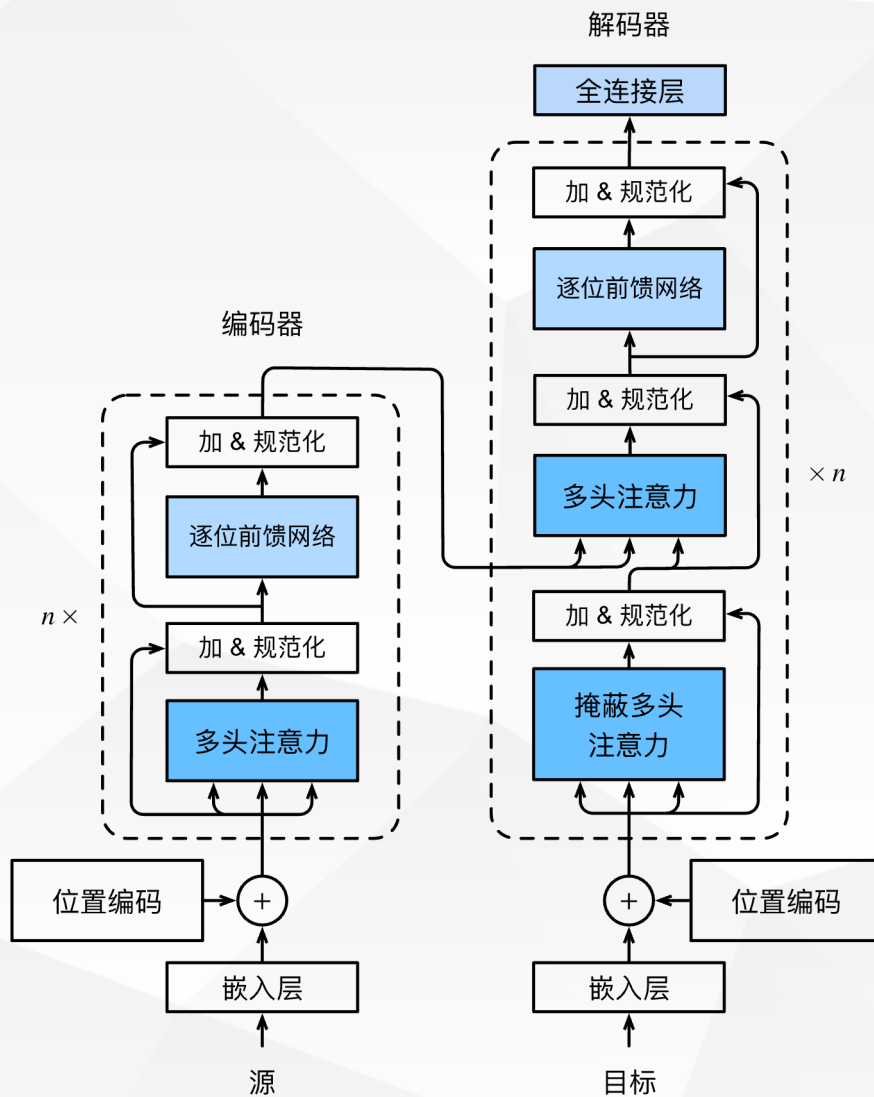
② 位置编码

③ 残差连接

④ 层规范化

⑤ 掩蔽注意力机制

⑥ 代码实现





① 在 Transformer 的基础上，自选问题进行研究。

- 将Transformer的结构中，是先进行残差连接再进行LayerNorm操作的。尝试替换顺序，观察网络性能变化结果，并结合相关文献分析原因。
- Transformer最初设计是被用于自然语言处理任务中的。调研相关文献，研究将Transformer网络结构应用到计算机视觉领域的方案，分析其相对于CNN的优劣。
- 其他，合理即可。

② 作业提交：实验报告+源代码（如有）

③ 截止日期：2023-10-27 23:59