



智能科学与技术学院



南京大學
NANJING UNIVERSITY

自然语言处理

6. 卷积神经网络

虞剑飞

南京大学智能科学与技术学院

2025.4.2

本章内容

- 概述
- 卷积
- 二维卷积
- 卷积神经网络
 - 卷积层
 - 汇聚层
 - 典型结构
- 文本卷积神经网络
- 卷积神经网络的应用

本章内容

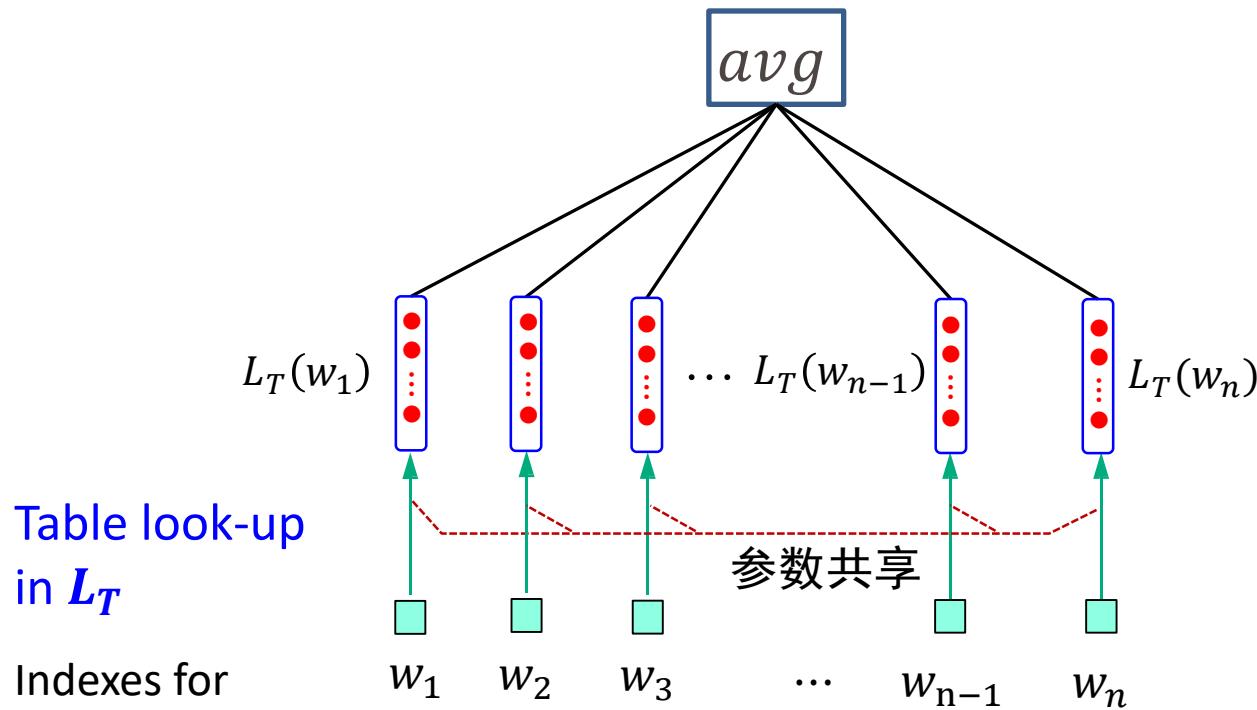
- 概述
- 卷积
- 二维卷积
- 卷积神经网络
 - 卷积层
 - 汇聚层
 - 典型结构
- 文本卷积神经网络
- 卷积神经网络的应用

句子的表示方法

- 词向量的平均

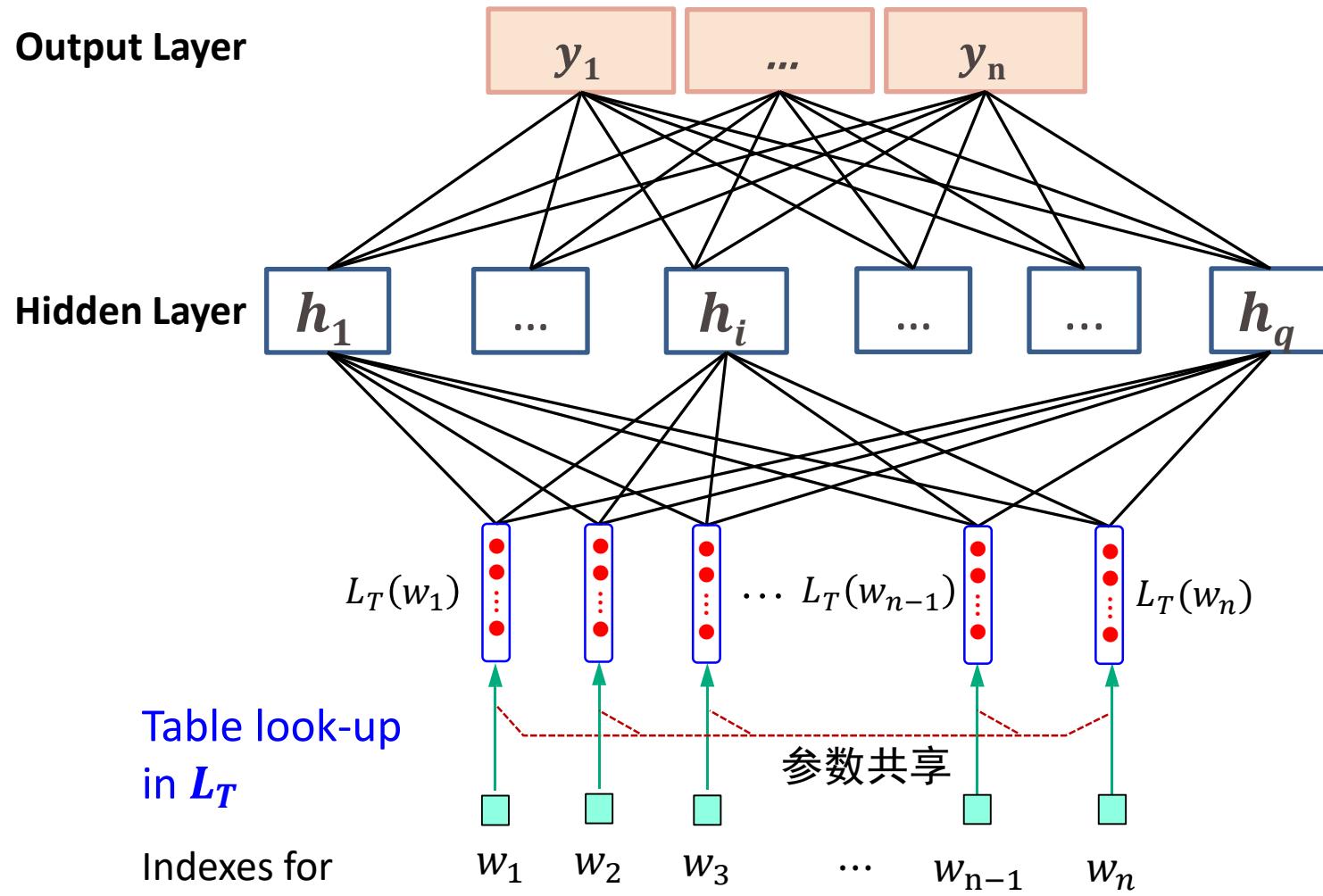
- 问题

- 无法考虑词序



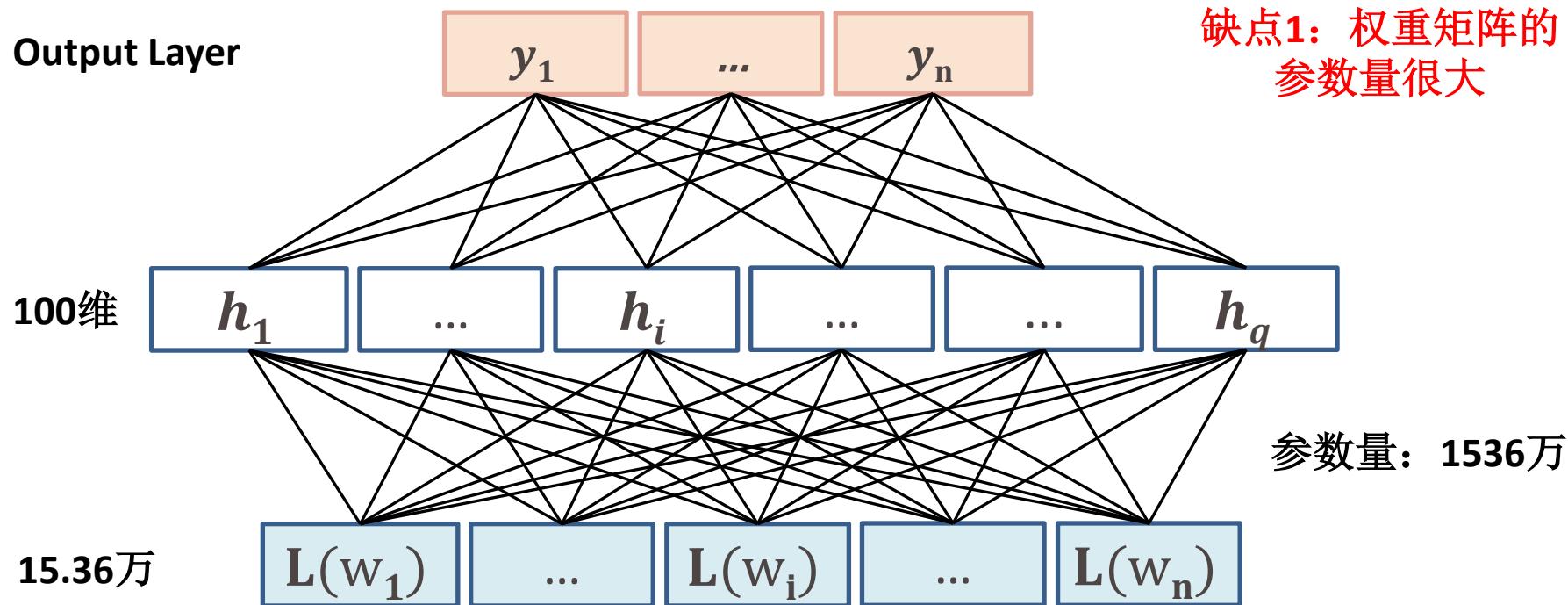
三层前向神经网络

- 基于词向量的前向神经网络模型



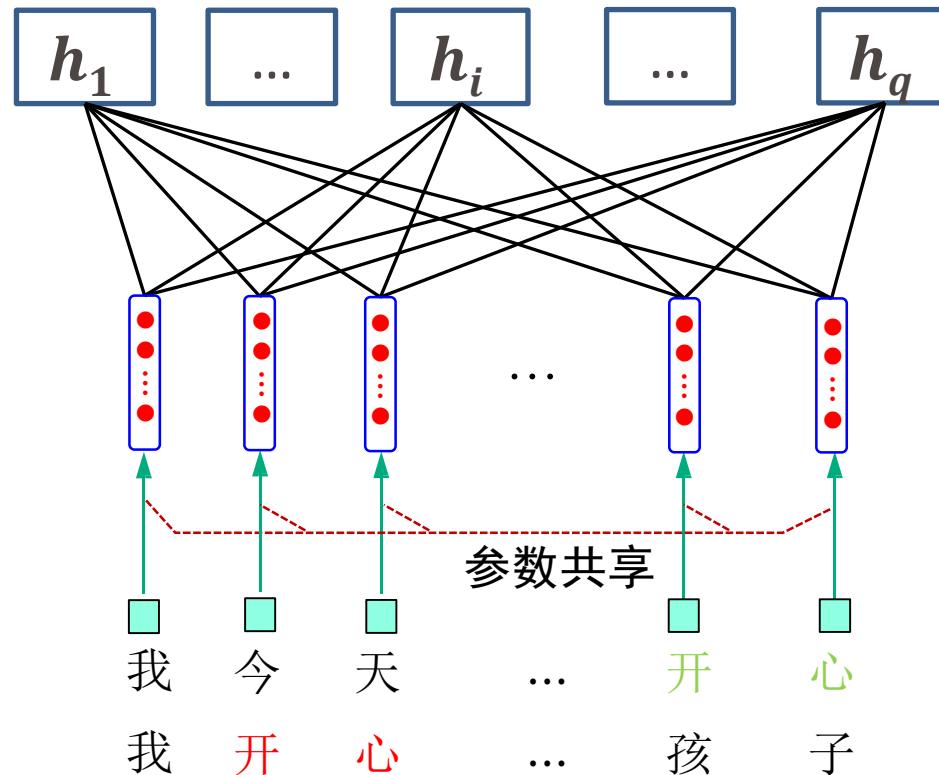
三层前向神经网络

- 情感分类应用：预测文本的情感倾向（正面vs中立vs负面）
 - 使用词向量作为输入（300维）
 - 文档最大长度512（长度不足时padding）
 - 假设隐藏层神经元个数设为100个



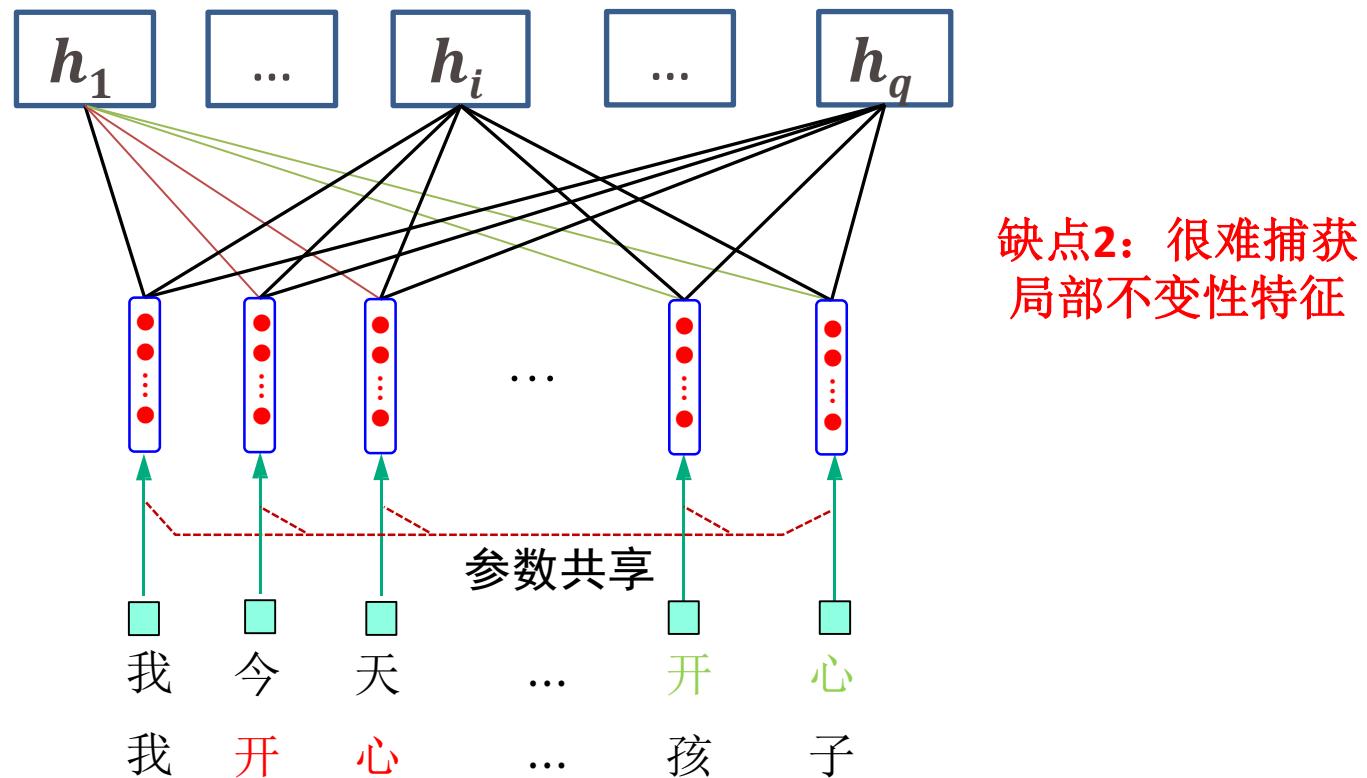
三层前向神经网络

- 局部不变性特征
 - 自然语言中的短语一般而言具有**局部不变性特征**
 - 出现在句中不同位置一般不影响其语义信息



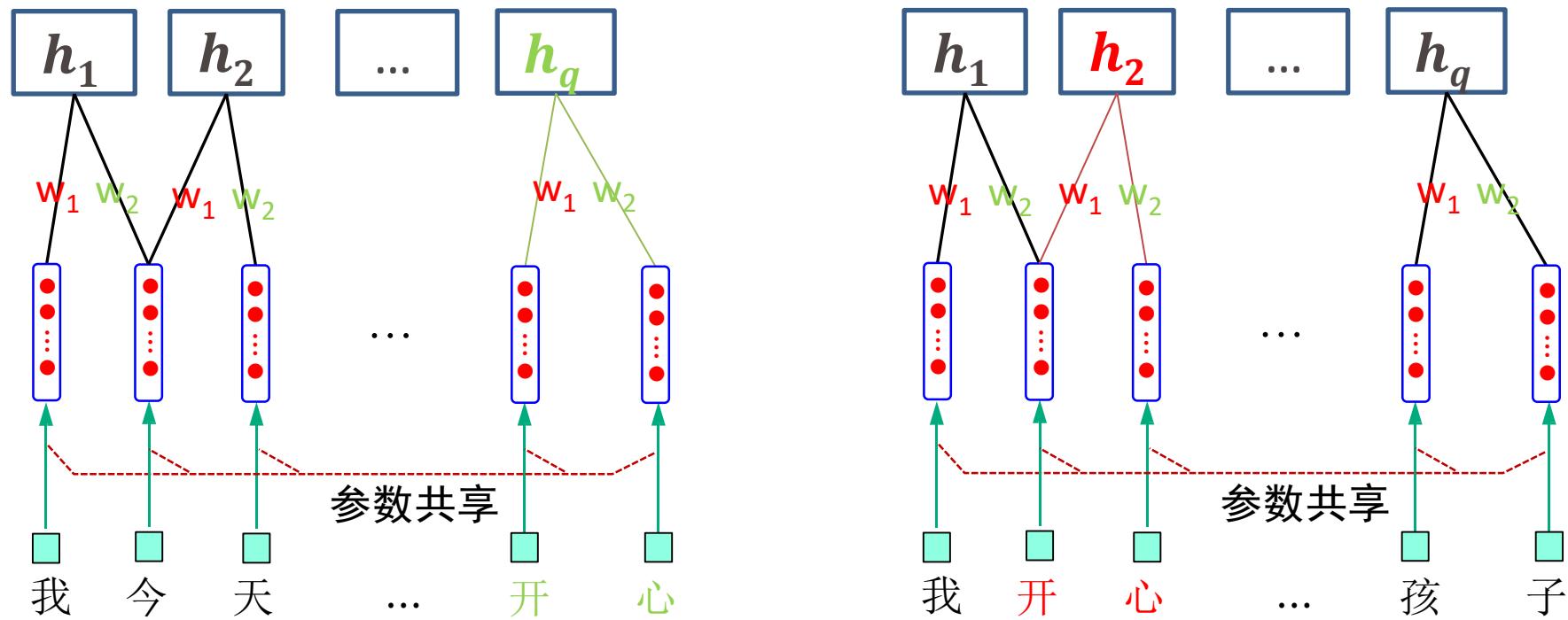
三层前向神经网络

- 局部不变性特征
 - 自然语言中的短语一般而言具有**局部不变性特征**
 - 出现在句中不同位置一般不影响其语义信息



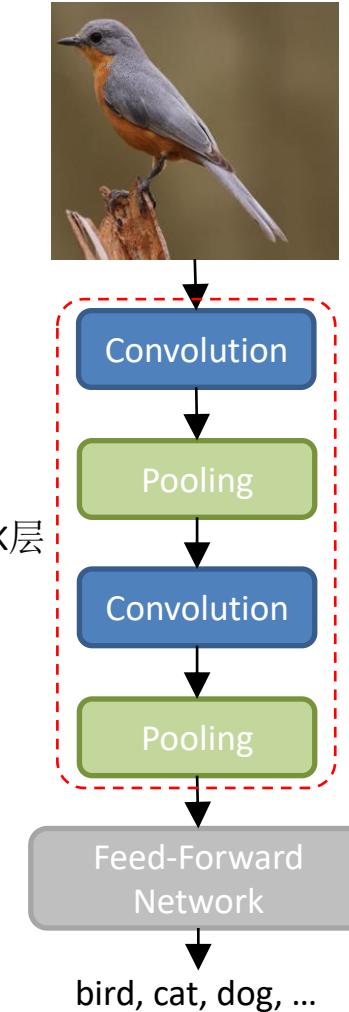
三层前向神经网络

- 每个隐藏层神经元只与局部输入相连
 - 2-gram



卷积神经网络

- 卷积神经网络 (Convolutional Neural Networks, CNN)
 - 一种前向神经网络
 - 包含若干卷积层和汇聚层（池化层）
 - 受生物学上感受野 (Receptive Field) 的机制而提出的
 - 在视觉神经系统中，一个神经元的感受野是指视网膜上的特定区域，只有这个区域内的刺激才能够激活该神经元。
- 结构特性
 - 权重共享
 - 解决FFN中参数量过大的问题
 - 局部连接
 - 解决FFN很难捕获局部不变性特征的缺陷
 - 空间或时间上的次采样



本章内容

- 概述
- 卷积
- 二维卷积
- 卷积神经网络
 - 卷积层
 - 汇聚层
 - 典型结构
- 文本卷积神经网络
- 卷积神经网络的应用

卷积

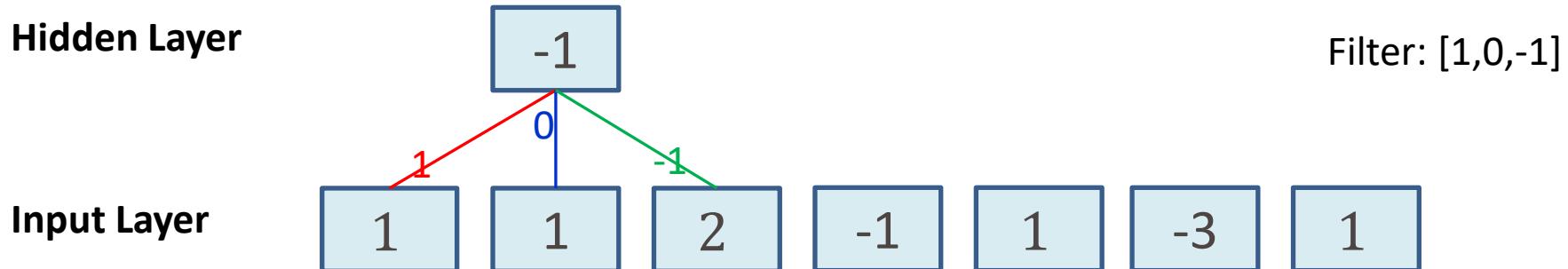
- 卷积的定义 (Convolution)

- 给定一个输入信号序列: $x = [x_1, x_2, x_3, \dots, x_d]$
- 滤波器filter (或者卷积核convolutional kernel) : $w = [w_1, w_2, \dots, w_k]$
- 在 t 时刻的卷积结果为第 t 个时刻以及前 $K-1$ 个时刻信息的叠加:

$$y_t = \sum_{u=1}^K w_{K-u+1} x_{t-u+1}$$

- 假设 $w = [w_1, w_2, w_3]$, 卷积的输出为:

$$y_t = w_1 x_{t-2} + w_2 x_{t-1} + w_3 x_t$$



卷积

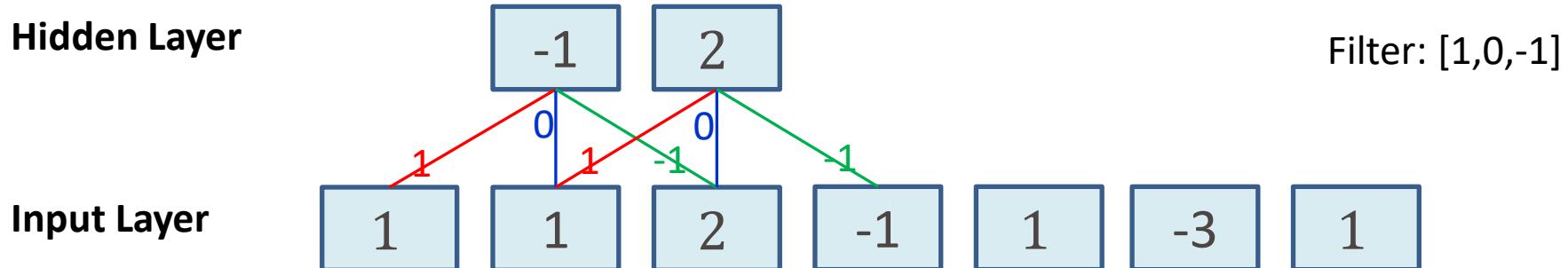
- 卷积的定义 (Convolution)

- 给定一个输入信号序列: $x = [x_1, x_2, x_3, \dots, x_d]$
- 滤波器filter (或者卷积核convolutional kernel) : $w = [w_1, w_2, \dots, w_k]$
- 在 t 时刻的卷积结果为第 t 个时刻以及前 $K-1$ 个时刻信息的叠加:

$$y_t = \sum_{u=1}^K w_{K-u+1} x_{t-u+1}$$

- 假设 $w = [w_1, w_2, w_3]$, 卷积的输出为:

$$y_t = w_1 x_{t-2} + w_2 x_{t-1} + w_3 x_t$$



卷积

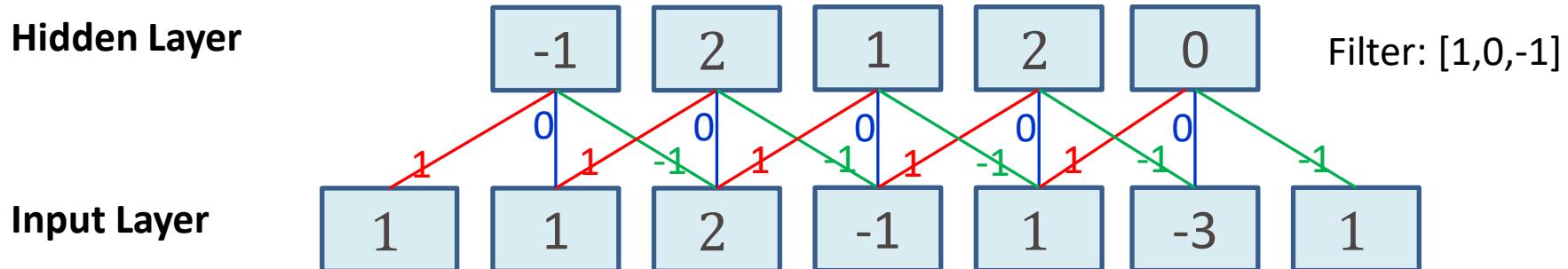
- 卷积的定义 (Convolution)

- 给定一个输入信号序列: $x = [x_1, x_2, x_3, \dots, x_d]$
- 滤波器filter (或者卷积核convolutional kernel) : $w = [w_1, w_2, \dots, w_k]$
- 在 t 时刻的卷积结果为第 t 个时刻以及前 $K-1$ 个时刻信息的叠加:

$$y_t = \sum_{u=1}^K w_{K-u+1} x_{t-u+1}$$

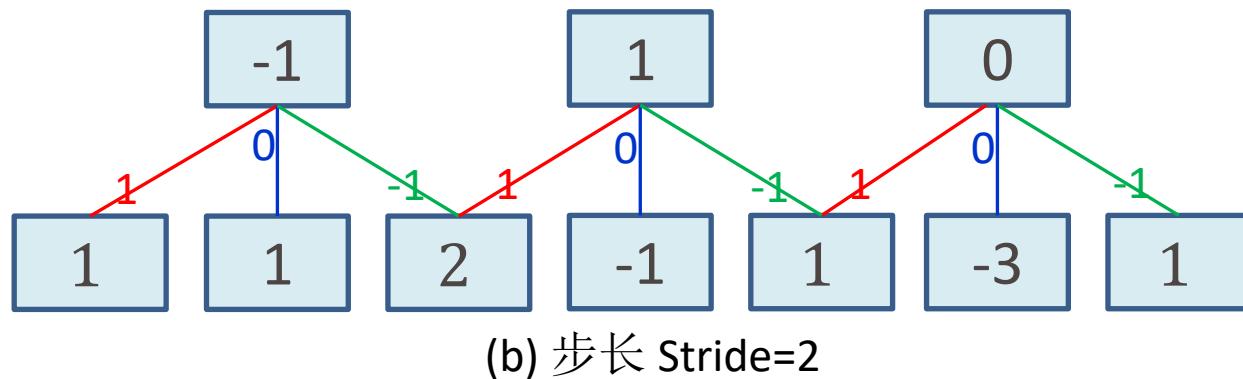
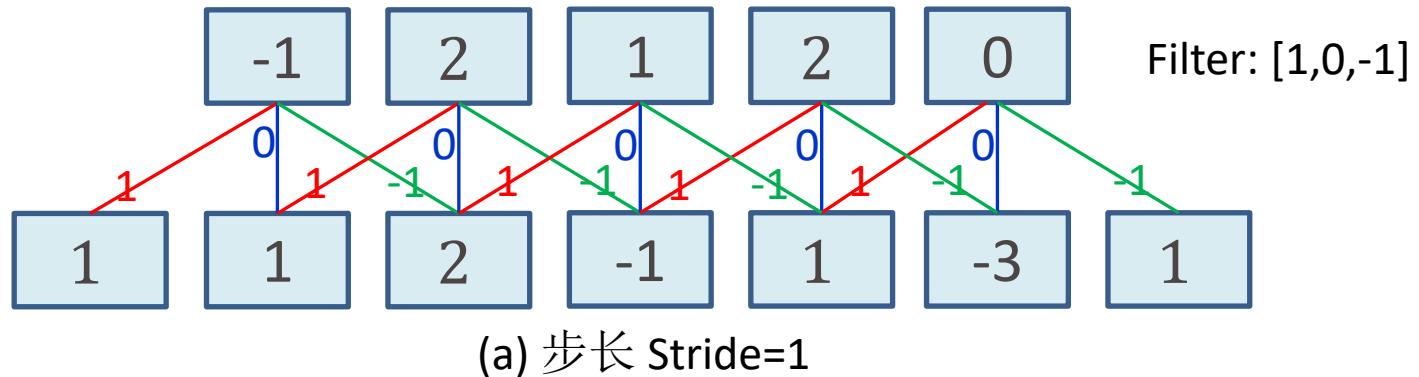
- 假设 $w = [w_1, w_2, w_3]$, 卷积的输出为:

$$y_t = w_1 x_{t-2} + w_2 x_{t-1} + w_3 x_t$$



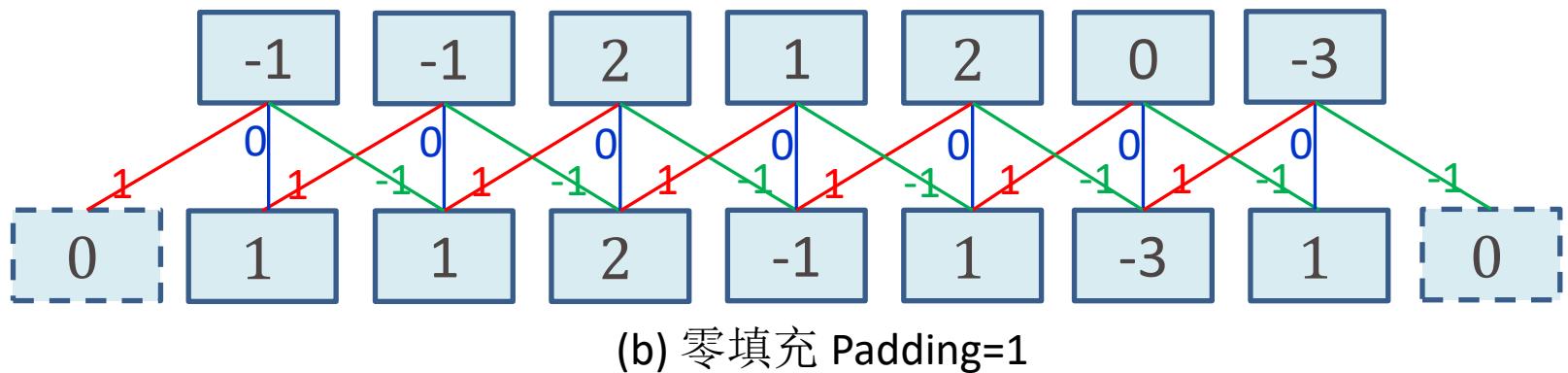
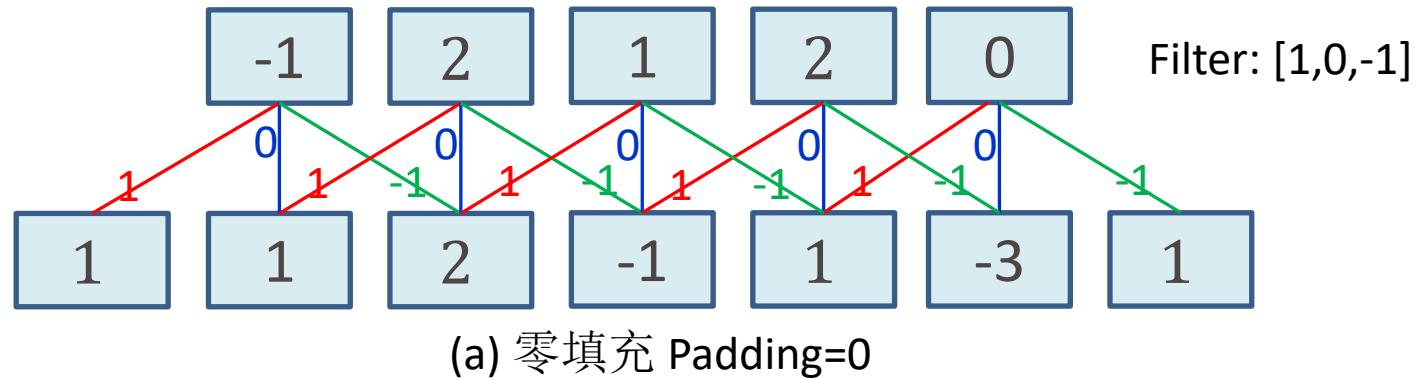
卷积

- 卷积的扩展
 - 步长 (Stride)



卷积

- 卷积的扩展
 - 填充 (Padding)



卷积

- 卷积类型
 - 输入长度为 M 的序列卷积的结果长度为
 - $L = \lfloor (M + 2 * P - K) / S + 1 \rfloor$
 - 卷积的结果按输出长度不同可以分为三类：
 - 窄卷积：步长 $S = 1$ ，两端不补零 $P = 0$ ，卷积后输出长度为 $M - K + 1$
 - 等宽卷积：步长 $S = 1$ ，两端补零 $P = (K - 1)/2$ ，卷积后输出长度 M
 - 宽卷积：步长 $S = 1$ ，两端补零 $P = K - 1$ ，卷积后输出长度 $M + K - 1$
 - 在早期的文献中，卷积一般默认为窄卷积。
 - 而目前的文献中，卷积一般默认为等宽卷积。

本章内容

- 概述
- 卷积
- 二维卷积
- 卷积神经网络
 - 卷积层
 - 汇聚层
 - 典型结构
- 文本卷积神经网络
- 卷积神经网络的应用

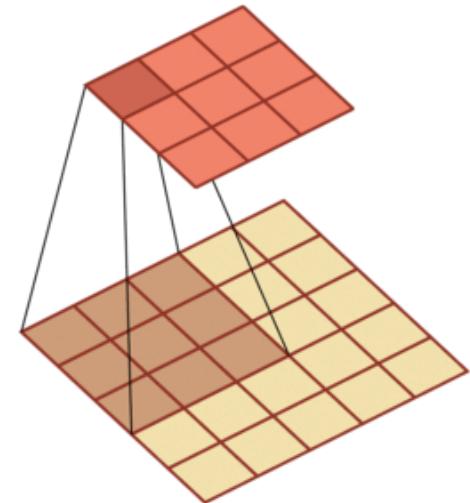
二维卷积

- 在图像处理中，图像是以二维矩阵的形式输入到神经网络中，因此我们需要二维卷积。
- 二维卷积
 - 给定一个输入: x
 - 滤波器filter: w
 - 二维卷积结果:

$$y_t = \sum_{u=1}^K w_{K-u+1} x_{t-u+1}$$

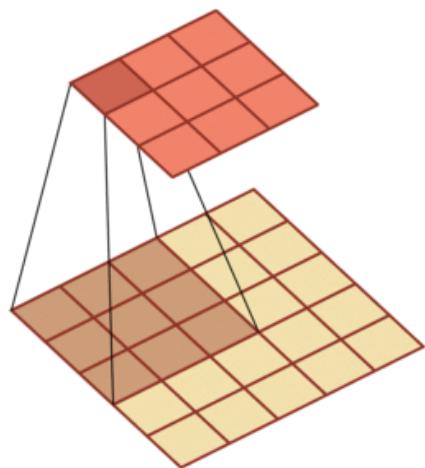
↓

$$y_{ij} = \sum_{u=1}^U \sum_{v=1}^V w_{U-u+1, V-v+1} x_{i-u+1, j-v+1}$$

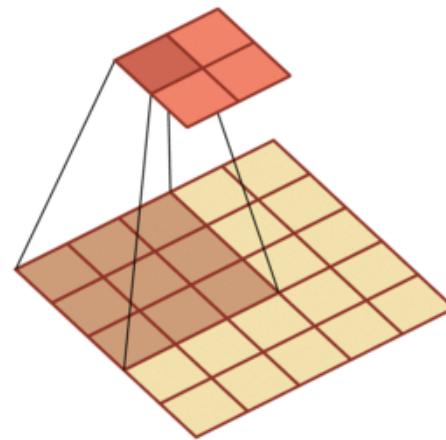


二维卷积

- 二维卷积的扩展
 - 步长 (Stride)



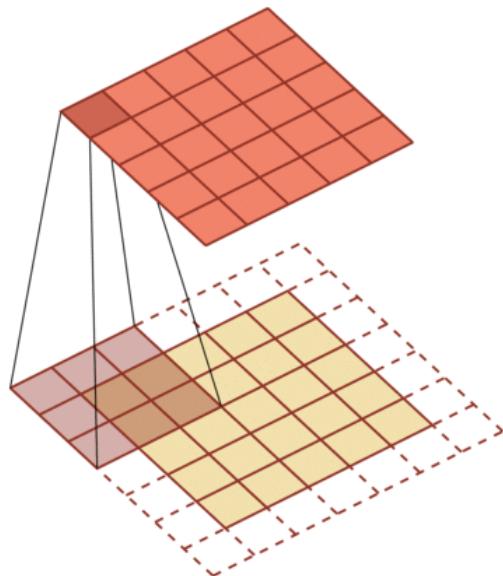
步长1，零填充0



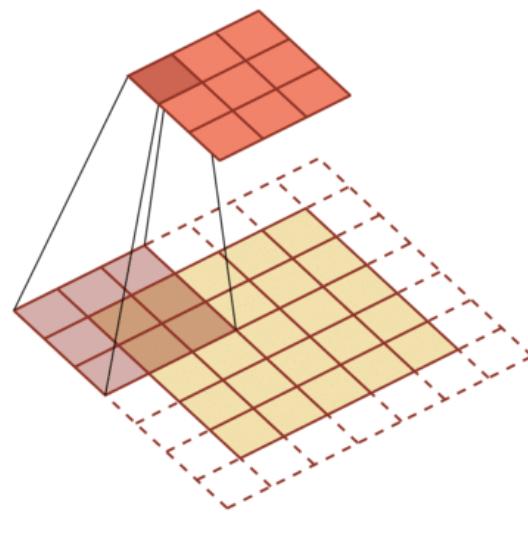
步长2，零填充0

二维卷积

- 二维卷积的扩展
 - 填充 (Padding)



步长1，零填充1



步长2，零填充1

二维卷积

- 课堂练习

3	0	1	2	7	4
1	5	8	9	3	1
2	7	2	5	1	3
0	1	3	1	7	8
4	2	1	6	2	8
2	4	5	2	3	9

6×6 (Gray image: $6 \times 6 \times 1$)

$$\begin{array}{|c|c|c|} \hline 1 & 0 & -1 \\ \hline 1 & 0 & -1 \\ \hline 1 & 0 & -1 \\ \hline \end{array} * \begin{array}{|c|c|c|} \hline = ? \\ \hline \end{array}$$

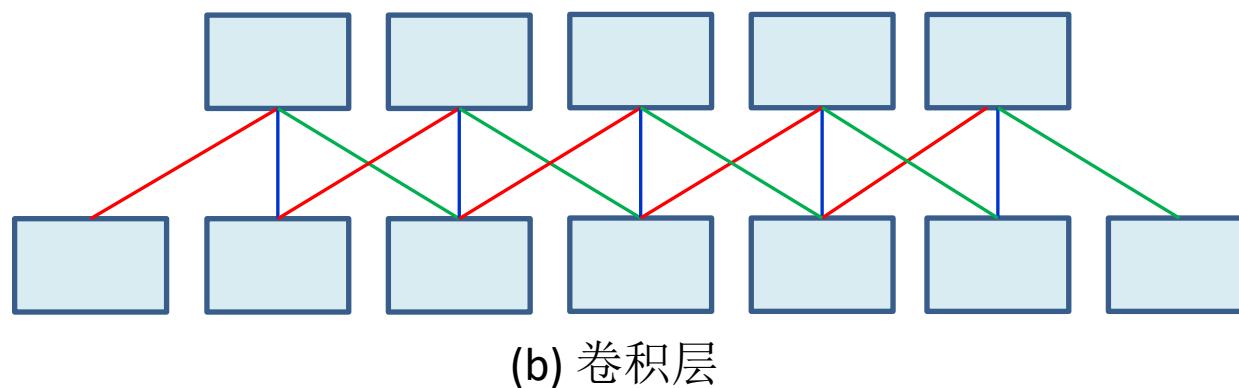
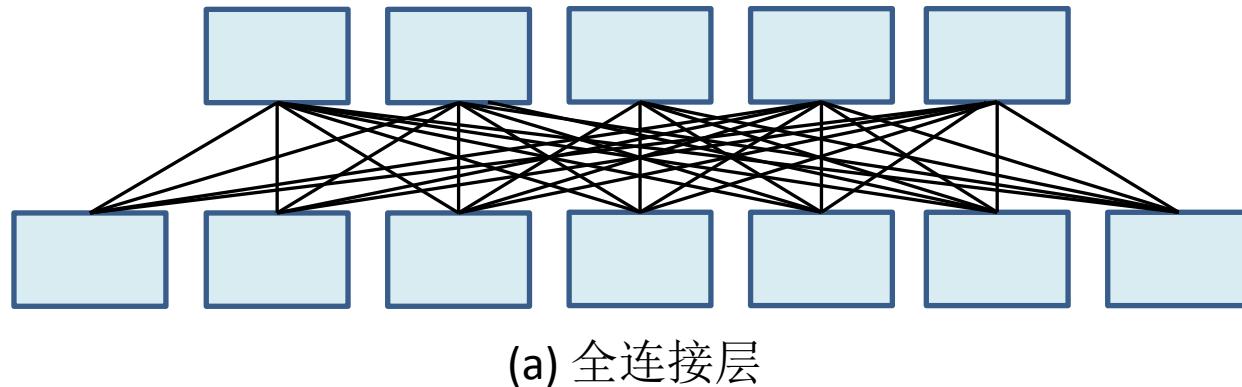
**3 x 3 filter/kernel
Padding=0
Stride=1**

本章内容

- 概述
- 卷积
- 二维卷积
- 卷积神经网络
 - 卷积层
 - 汇聚层
 - 典型结构
- 文本卷积神经网络
- 卷积神经网络的应用

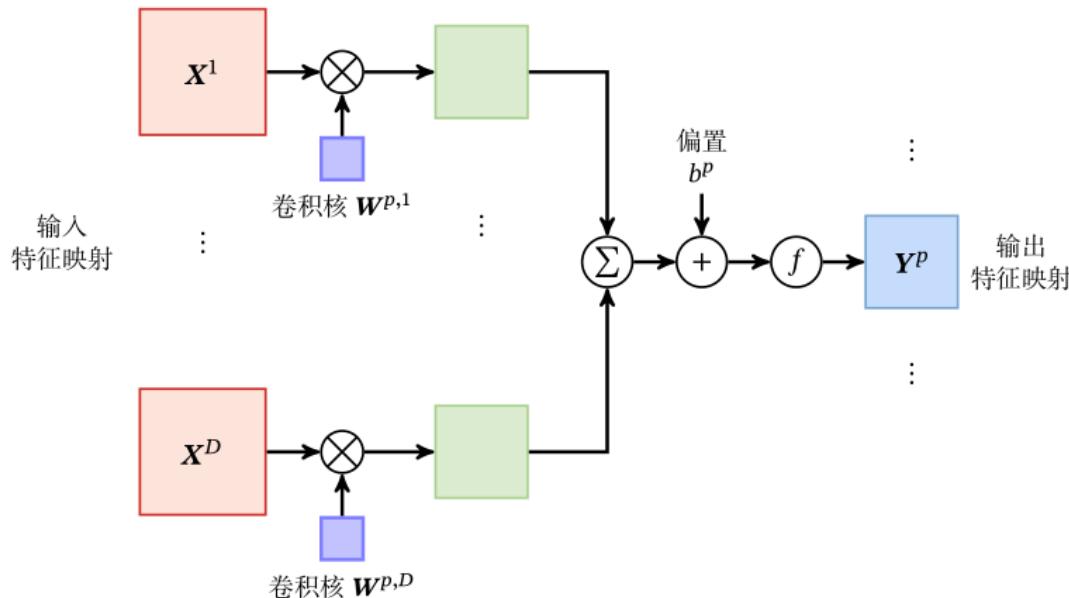
卷积神经网络

- 将全连接层替换为卷积层



卷积神经网络

- 卷积核作为一个特征提取器
 - 如何增强卷积层的能力?
 - 引入多个卷积核



$$\mathbf{Z}^p = \mathbf{W}^p \otimes \mathbf{X} + b^p = \sum_{d=1}^D \mathbf{W}^{pd} \otimes \mathbf{X}^d + b^p \quad \mathbf{Y}^p = f(\mathbf{Z}^p)$$

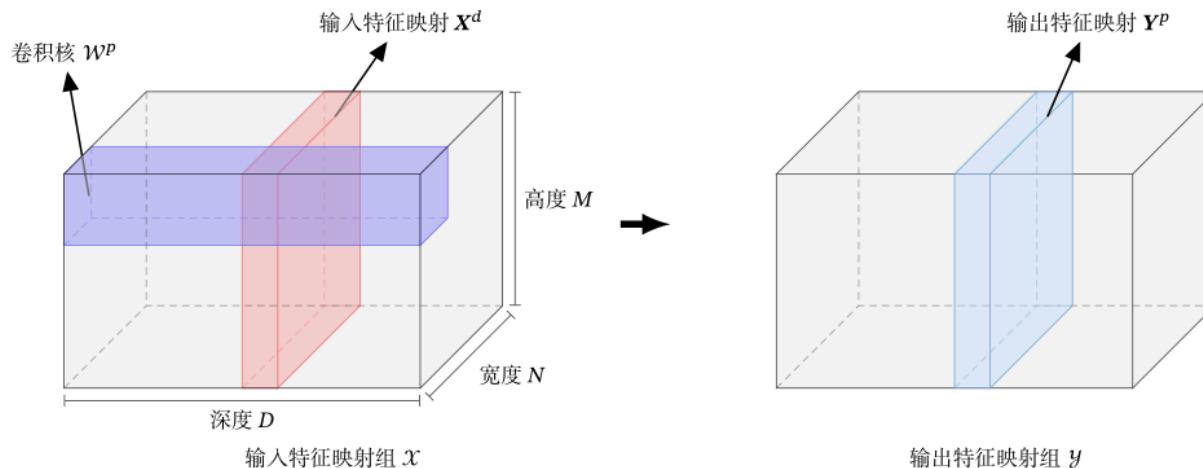
卷积神经网络



步长2
filter 3*3*3
filter个数2
零填充 1

卷积神经网络

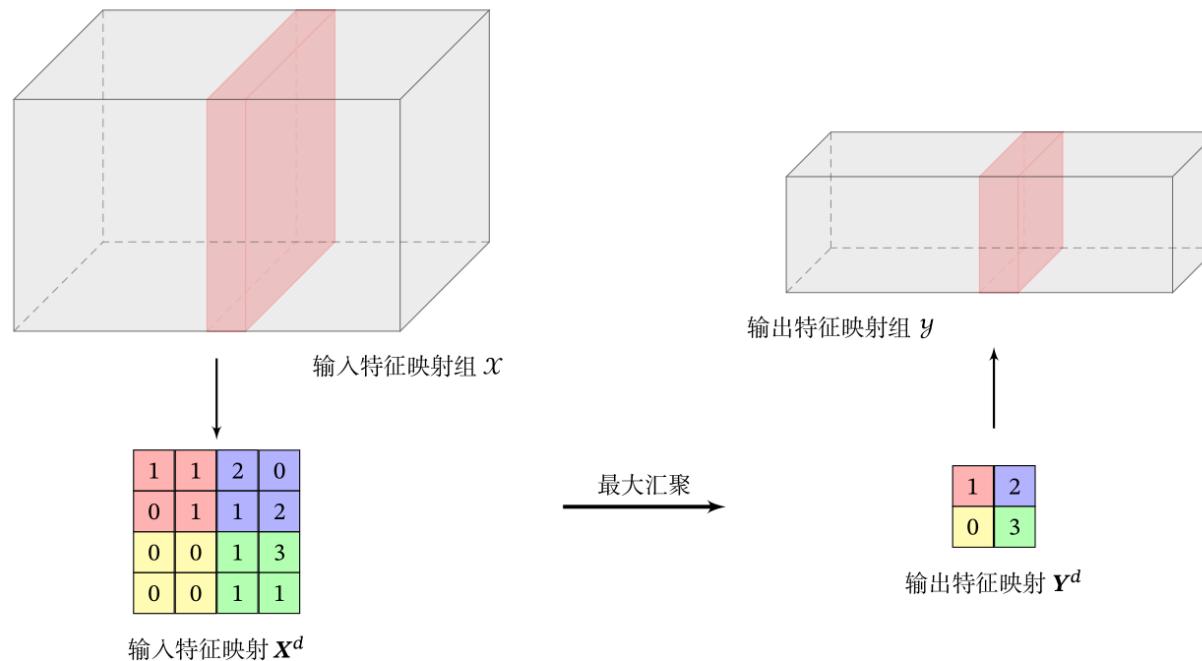
- 卷积层
 - 典型的卷积层为3维结构



$$\mathbf{Z}^p = \mathbf{W}^p \otimes \mathbf{X} + b^p = \sum_{d=1}^D \mathbf{W}^{pd} \otimes \mathbf{X}^d + b^p \quad \mathbf{Y}^p = f(\mathbf{Z}^p)$$

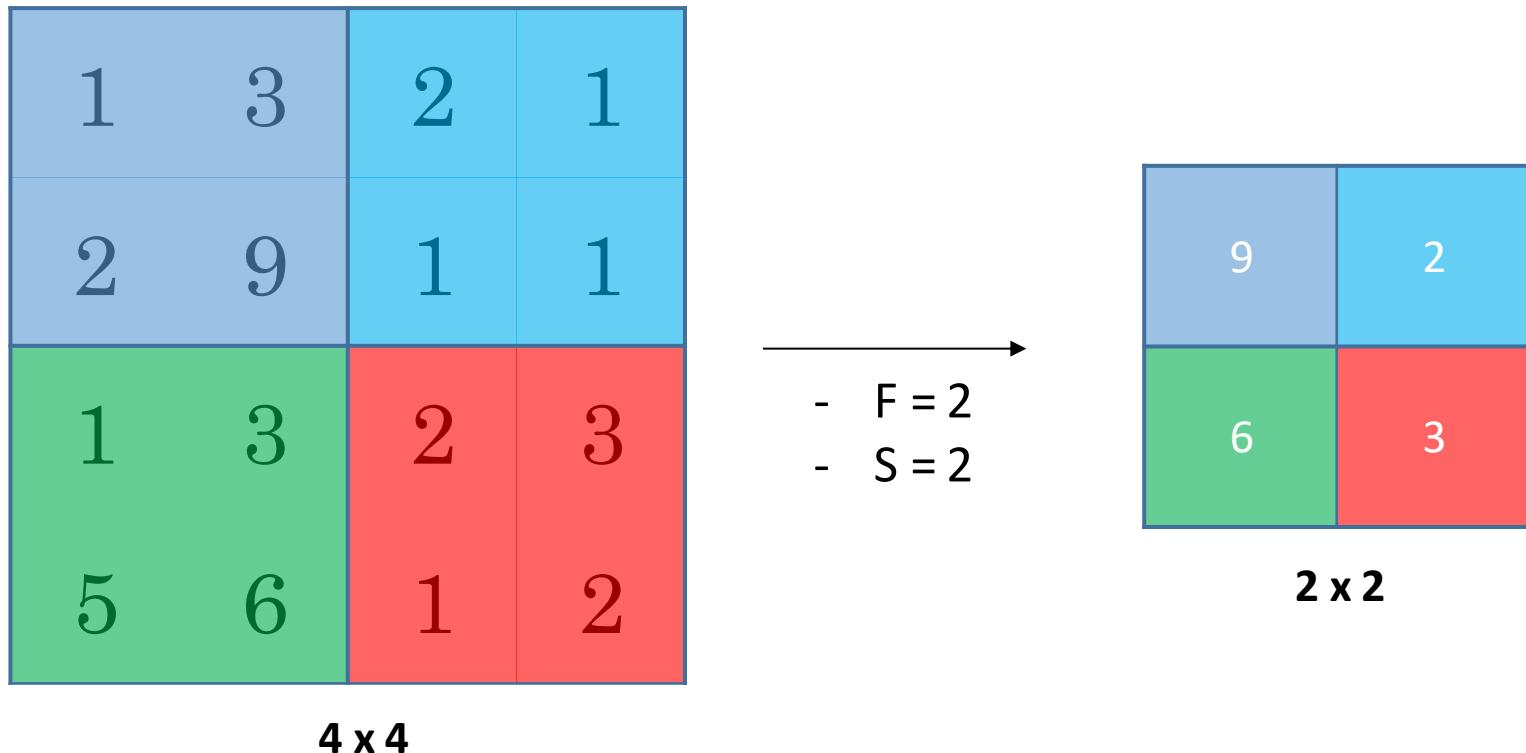
卷积神经网络

- 汇聚层 (也叫池化层, Pooling Layer)
 - 卷积层虽然可以显著减少连接的个数，但是每一个特征映射的神经元个数并没有显著减少。



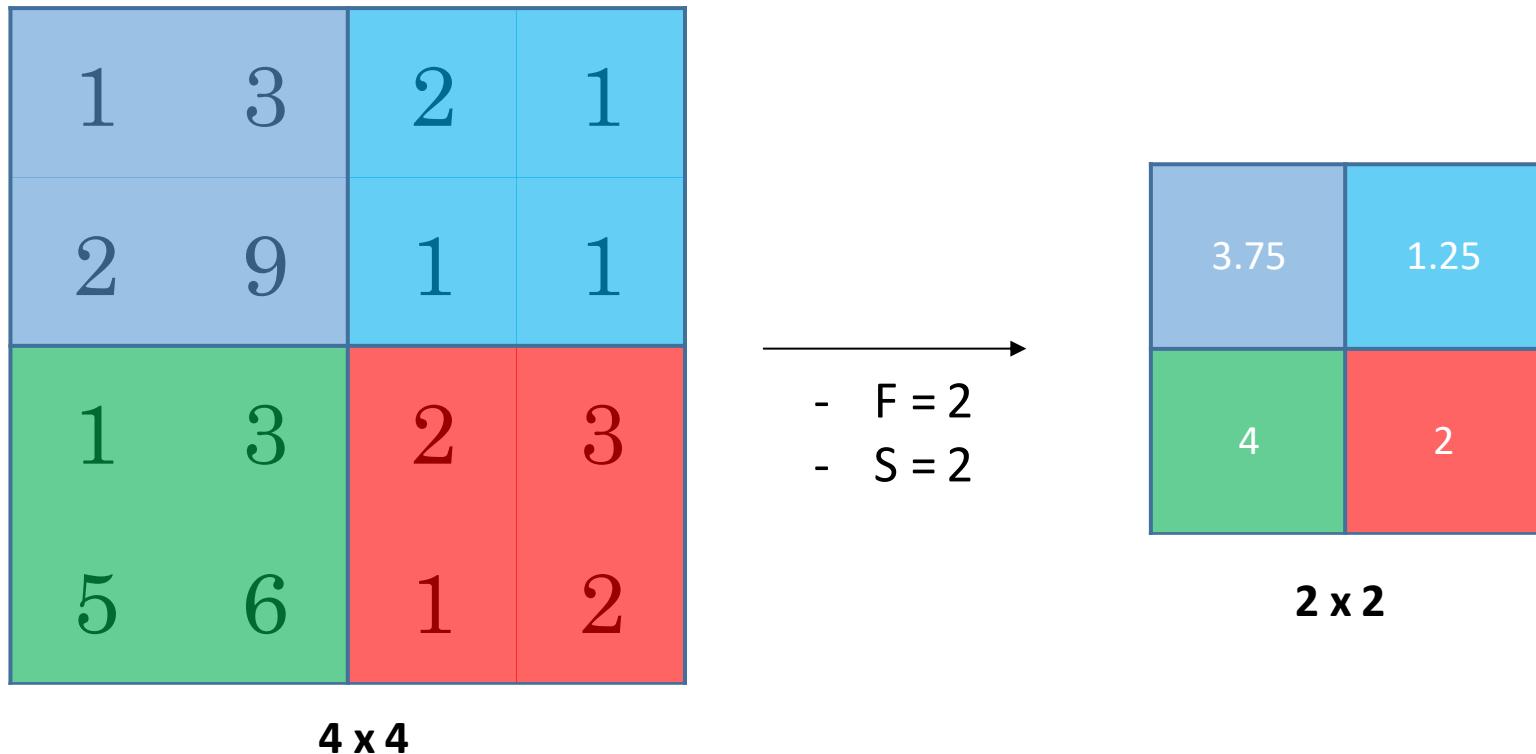
卷积神经网络

- 汇聚层 (也叫池化层, Pooling Layer)
 - 最大汇聚



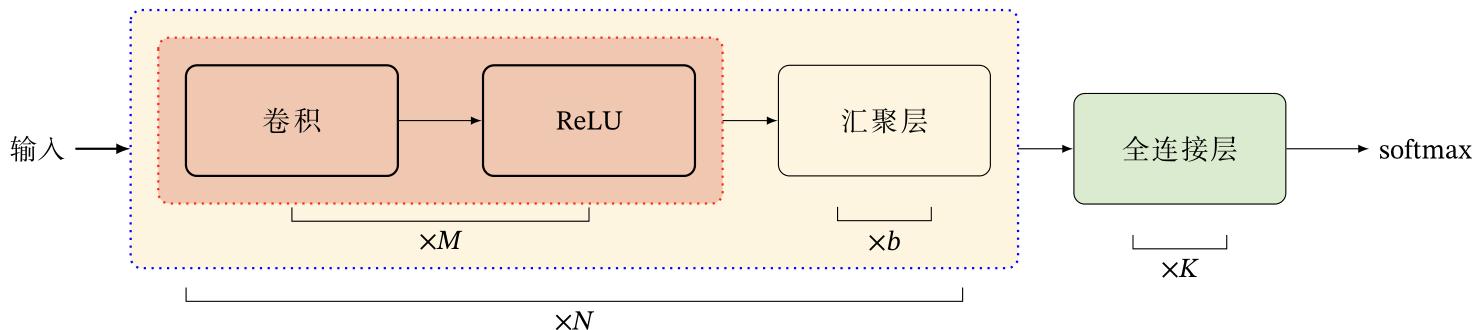
卷积神经网络

- 汇聚层 (也叫池化层, Pooling Layer)
 - 平均汇聚



卷积神经网络

- 卷积网络是由卷积层、汇聚层、全连接层交叉堆叠而成。



- 一个卷积块为连续M个卷积层和b个汇聚层（M通常设置为2~5，b为0或1）。一个卷积网络中可以堆叠N个连续的卷积块，然后在接着K个全连接层（N的取值区间比较大，比如1~100或者更大；K一般为0~2）。

- 典型结构

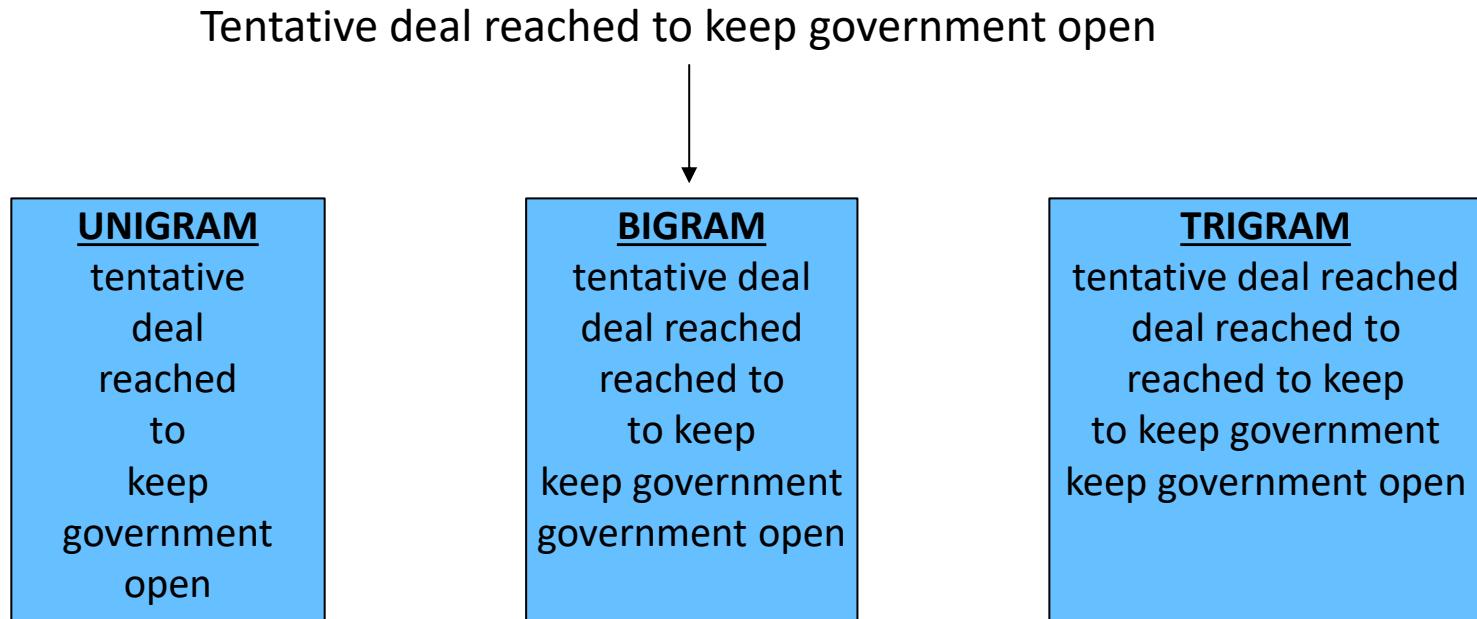
- 趋向于小卷积、大深度
- 趋向于全卷积

本章内容

- 概述
- 卷积
- 二维卷积
- 卷积神经网络
 - 卷积层
 - 汇聚层
 - 典型结构
- 文本卷积神经网络
- 卷积神经网络的应用

文本卷积神经网络

- N-gram特征
 - 以文本分类任务为例



如何用卷积操作来实现？

文本卷积神经网络

- 词向量
 - 通过一个数字组成的d维向量来表示一个单词
 - 示例

tentative	0.2	0.1	-0.3	0.4
deal	0.5	0.2	-0.3	-0.1
reached	-0.1	-0.3	-0.2	0.4
to	0.3	-0.3	0.1	0.1
keep	0.2	-0.3	0.4	0.2
government	0.1	0.2	-0.1	-0.1
open	-0.4	-0.4	0.2	0.3

<https://code.google.com/archive/p/word2vec/>

<https://nlp.stanford.edu/projects/glove/>

文本卷积神经网络

- 一维卷积
 - 课堂练习

tentative	0.2	0.1	-0.3	0.4
deal	0.5	0.2	-0.3	-0.1
reached	-0.1	-0.3	-0.2	0.4
to	0.3	-0.3	0.1	0.1
keep	0.2	-0.3	0.4	0.2
government	0.1	0.2	-0.1	-0.1
open	-0.4	-0.4	0.2	0.3

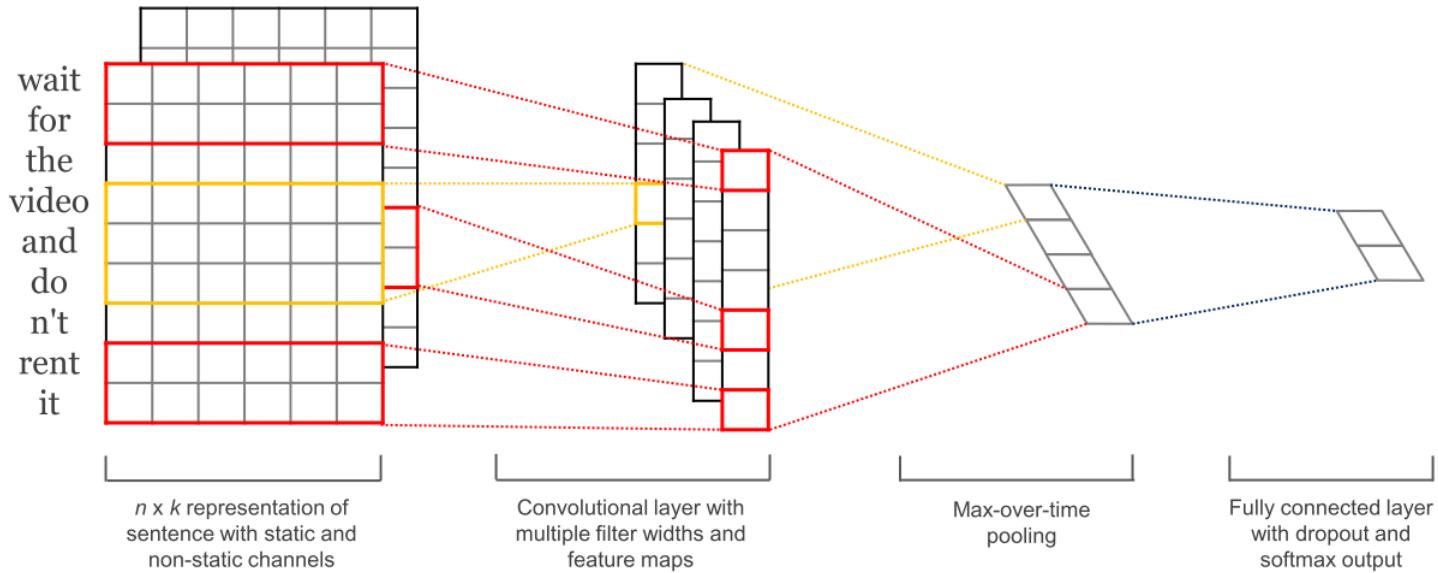
$N \times 4$

$$\begin{matrix} & 3 & 1 & 2 & -3 \\ * & -1 & 2 & 1 & -3 \\ & 1 & 1 & -1 & 1 \end{matrix} = ?$$

3×4

文本卷积神经网络

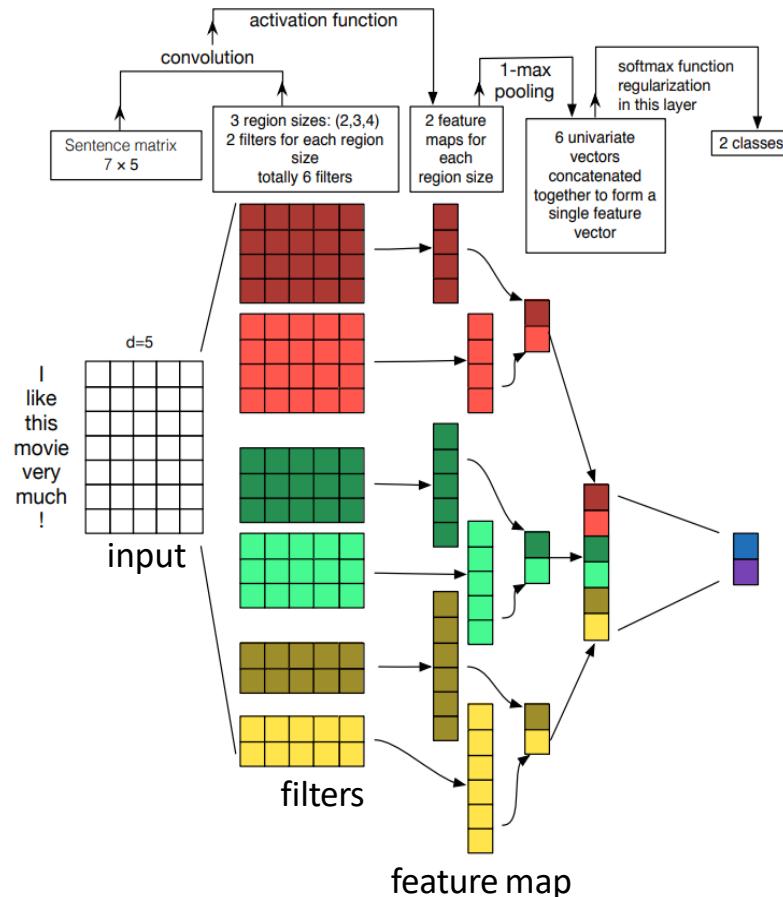
- 基于卷积模型的句子表示



Y. Kim. “Convolutional neural networks for sentence classification”. In EMNLP 2014.

文本卷积神经网络

- 基于卷积模型的句子表示



Zhang, Ye, and Byron Wallace. "A sensitivity analysis of (and practitioners' guide to) convolutional neural networks for sentence classification." arXiv preprint arXiv:1510.03820 (2015).

文本卷积神经网络

- 代码实现

batch_size = 1

word_embed_size = 4

seq_len = 7

tentative	deal	reached	to	keep	government	open
0.2	0.5	-0.1	0.3	0.2	0.1	-0.4
0.1	0.2	-0.3	-0.3	-0.3	0.2	-0.4
-0.3	-0.3	-0.2	0.1	0.4	-0.1	0.2
0.4	-0.1	0.4	0.1	0.2	-0.1	0.3

Tri-gram filter: 4x3

```
input = torch.randn(batch_size, seq_len, word_embed_size) # input: (1, 7, 4)
conv1 = nn.Conv1d(in_channels=word_embed_size, out_channels=3, # of filters
                 kernel_size=3) # can add padding=1
```

```
hidden1 = conv1(input) # hidden1: (1, 5, 3)
```

```
hidden2 = torch.max(hidden1, dim=1) # max pool; hidden2: (1, 3)
```

	t,d,r	d,r,t	r,t,k	t,k,g	k,g,o
out_channel 1	-1.0	-0.5	-3.6	-0.2	0.3
out_channel 2	1.6	-0.1	0.3	0.1	0.6
out_channel 3	-1.0	0.8	0.3	1.2	0.9

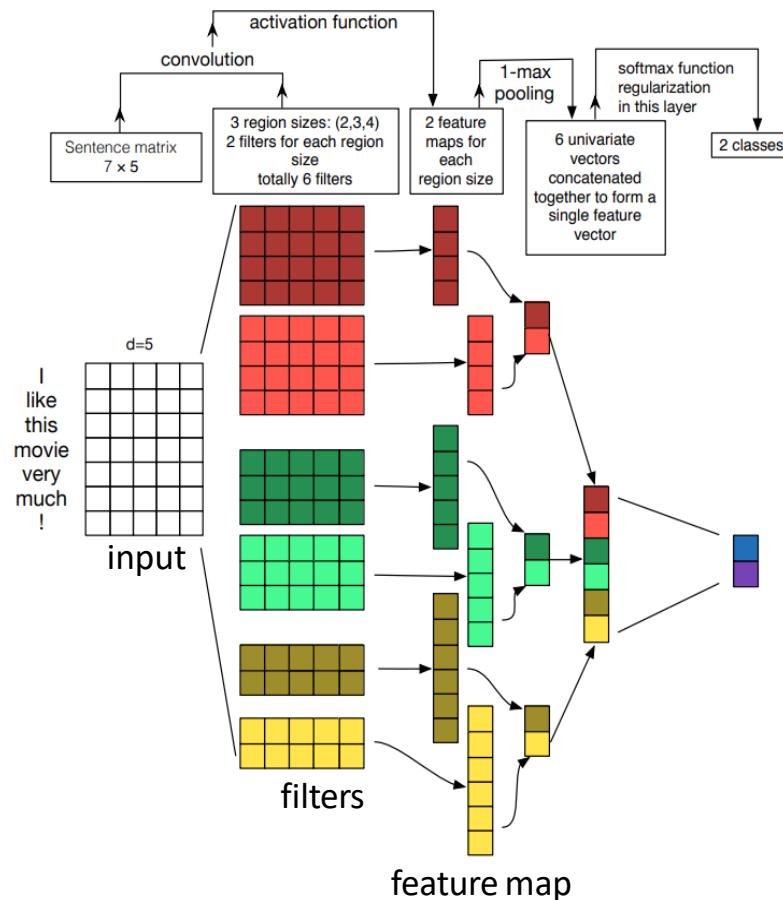


本章内容

- 概述
- 卷积
- 二维卷积
- 卷积神经网络
 - 卷积层
 - 汇聚层
 - 典型结构
- 文本卷积神经网络
- **卷积神经网络的应用**

文本分类

- 情感分类



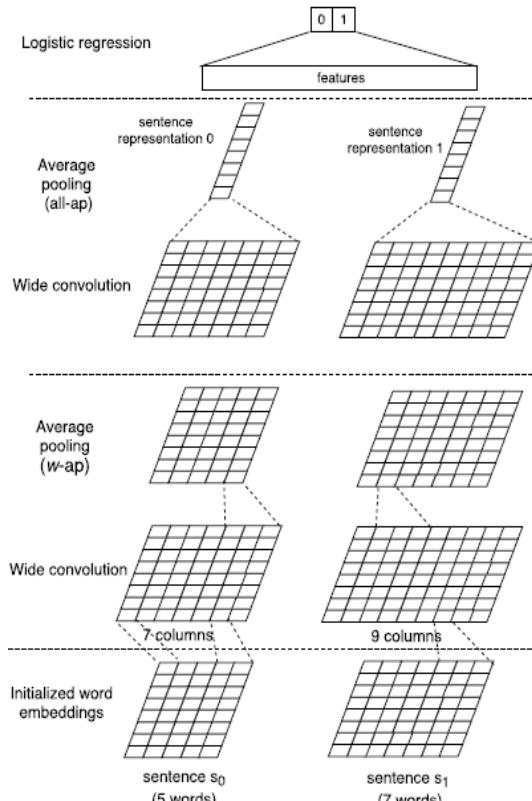
自然语言推断

- 自然语言推断 (Natural Language Inference)
 - 输入为两个句子
 - 前提 (Premise)
 - 假设 (Hypothesis)
 - 输出为两个句子之间的关系
 - 蕴含 (entailment)
 - 对立 (contradiction)
 - 中立 (neutral)

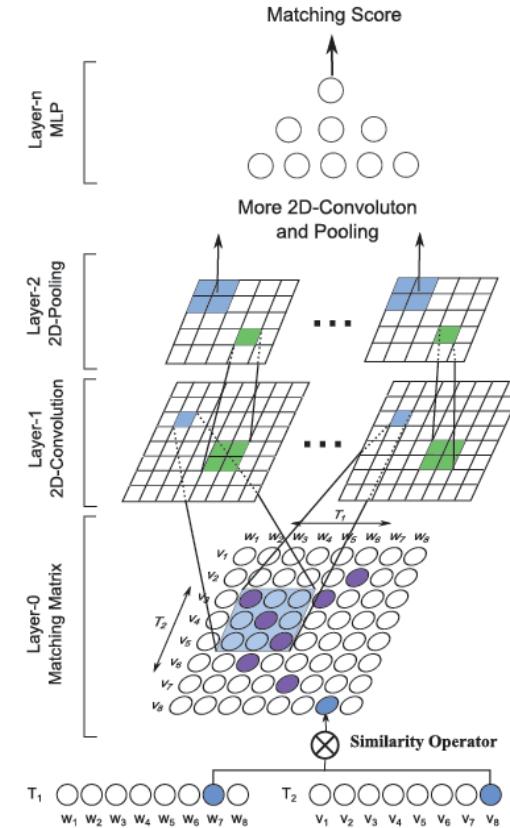
Text	Judgments	Hypothesis
A man inspects the uniform of a figure in some East Asian country.	contradiction C C C C C	The man is sleeping
An older and younger man smiling.	neutral N N E N N	Two men are smiling and laughing at the cats playing on the floor.
A black race car starts up in front of a crowd of people.	contradiction C C C C C	A man is driving down a lonely road.
A soccer game with multiple males playing.	entailment E E E E E	Some men are playing a sport.
A smiling costumed woman is holding an umbrella.	neutral N N E C N	A happy woman in a fairy costume holds an umbrella.

自然语言推断

- 典型模型



BCNN



Pyramid

课程作业

- 基于卷积神经网络的文本情感分类
 - 问题描述
 - 利用卷积神经网络，实现对一个电影评论数据集 (Rotten Tomatoes dataset) 的情感分类问题
 - 数据集
 - 链接： <https://nlp.stanford.edu/sentiment/> （原始数据）
 - 预处理后数据链接：https://pan.baidu.com/s/1EuG7OR5_xwU81-ZDrQ3cxg
 - 提取码：otz1
 - 句子级情感五分类（very negative, negative, neutral, positive, very positive）

		label
1	sentences	
2	['it', '"s", 'a', 'lovely', 'film', 'with', 'lovely', 'performances', 'by', 'buy', 'and', 'accorsi']	3
3	['no', 'one', 'goes', 'unindicted', 'here', '', 'which', 'is', 'probably', 'for', 'the', 'best']	2
4	['and', 'if', 'you', '"re", 'not', 'nearly', 'moved', 'to', 'tears', 'by', 'a', 'couple', 'of', 'scenes', '', 'you', '"ve", 'got', 'ice', 'v']	3
5	['a', 'warm', '', 'funny', '', 'engaging', 'film']	4
6	['uses', 'sharp', 'humor', 'and', 'insight', 'into', 'human', 'nature', 'to', 'examine', 'class', 'conflict', '', 'adolescent', 'ye']	4
7	['half', 'submarine', 'flick', '', 'half', 'ghost', 'story', '', 'all', 'in', 'one', 'criminally', 'neglected', 'film']	2
8	['entertains', 'by', 'providing', 'good', '', 'lively', 'company']	3

课程作业

- 基于卷积神经网络的文本情感分类
 - 问题描述
 - 利用卷积神经网络，实现对一个电影评论数据集 (Rotten Tomatoes dataset) 的情感分类问题
 - 数据集
 - 链接: <https://nlp.stanford.edu/sentiment/> (原始数据)
 - 预处理后数据链接: https://pan.baidu.com/s/1EuG7OR5_xwU81-ZDrQ3cxg
 - 提取码: otz1
 - 句子级情感五分类 (very negative, negative, neutral, positive, very positive)
 - 要求
 - Tensorflow或者Pytorch实现
 - 评价指标
 - 准确率
 - 参考论文
 - Convolutional Neural Networks for Sentence Classification <https://arxiv.org/abs/1408.5882>



欢迎提问！