

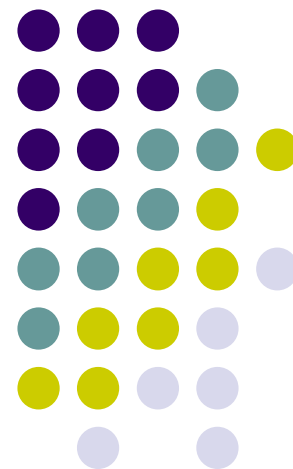
# 自然语言处理导论

---

张梅山

黑龙江大学

(<https://zhangmeishan.github.io/>)



# 主要内容



- 自然语言处理任务简介
- 文本分类
- 序列标注
- 结构学习
- 序列到序列

# 主要内容



- 自然语言处理任务简介
- 文本分类
- 序列标注
- 结构学习
- 序列到序列

# 基本印象



输入： 本次课将介绍自然语言处理相关的基本任务

输出：



# 基本印象



输入： 本次课将介绍自然语言处理相关的基本任务

输出：



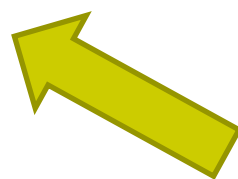
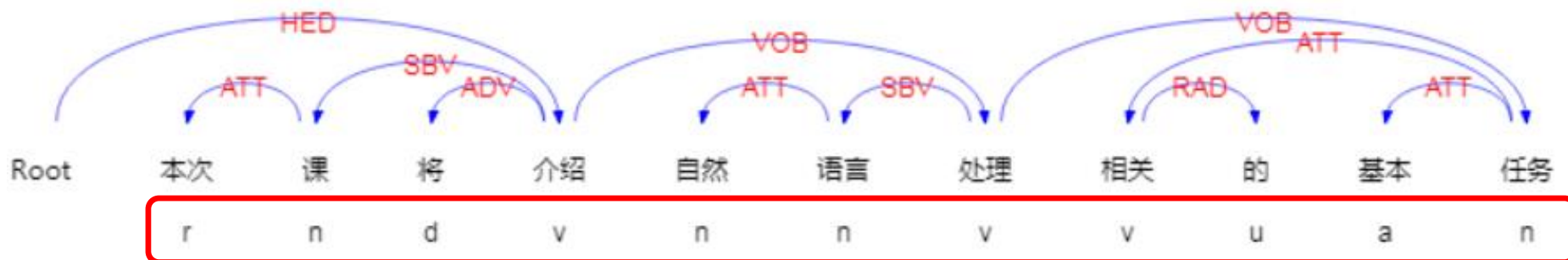
中文分词

# 基本印象



输入： 本次课将介绍自然语言处理相关的基本任务

输出：



词性/词类标注

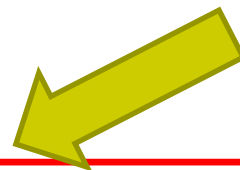
# 基本印象



输入： 本次课将介绍自然语言处理相关的基本任务

输出：

句法分析



# 自然语言处理任务介绍



## 基础分析

- 分词
- 词性标注
- 句法分析
- 语义分析
- 篇章分析



# 自然语言处理任务介绍



## 自然语言处理基本应用

- 情感分析
- 机器翻译
- 实体识别
- 阅读理解
- 智能对话

# 自然语言处理任务介绍



## 自然语言处理基本应用

- 情感分析
- 机器翻译
- 实体识别
- 阅读理解
- 智能对话

马拉多纳则斥责老对手英格兰在裁判判罚中获利

21/60

情感分析结果

负面  
73%



正面  
27%

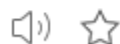
# 自然语言处理任务介绍



## 自然语言处理基本应用

- 情感分析
- 机器翻译
- 实体识别
- 阅读理解
- 智能对话

马拉多纳则斥责老对手英格兰在裁判判罚中获利



Maradona rebuked his old rival, England, for making profits in the referee's decision.



双语对照 ☐

# 自然语言处理任务介绍



## 自然语言处理基本应用

- 情感分析
- 机器翻译
- 实体识别
- 阅读理解
- 智能对话

马拉多纳则斥责老对手英格兰在裁判判罚中获利

人名

马拉多纳

地名

英格兰

# 自然语言处理任务介绍



## 自然语言处理基本应用

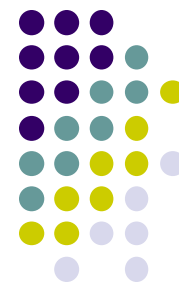
- 情感分析
- 机器翻译
- 实体识别
- 阅读理解
- 智能对话

马拉多纳则斥责老对手英格兰在裁判判罚中获利

问题：谁在裁判判罚中获利？

答：马拉多纳的老对手英格兰

# 自然语言处理任务介绍



## 自然语言处理基本应用

- 情感分析
- 机器翻译
- 实体识别
- 阅读理解
- 智能对话



上古的乒乓球



1小时前 - 马拉多纳近日在接受采访时,公开指责英格兰队在裁判判罚中获利,甚至直接指责国际足联为小偷,将这场比赛比喻为抢劫。



你这样有意思么



麻麻说捉弄人的都不是好孩子



真无聊,不理你了。。。

马拉多纳则斥责老对手英格兰在裁判判罚中获利

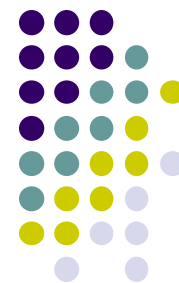
马拉多纳则斥责老对手英格兰在裁判判罚中获利

马拉多纳则斥责老对手英格兰在裁判判罚中获利

马拉多纳则斥责老对手英格兰在裁判判罚中获利

马拉多纳则斥责老对手英格兰在裁判判罚中获利

# 自然语言处理任务介绍



## 基础分析

- 分词
- 词性标注
- 句法分析
- 语义分析
- 篇章分析

## 自然语言处理基本应用

- 情感分析
- 机器翻译
- 实体识别
- 阅读理解
- 智能对话

目的：用计算机自动完成这一切。



## 一切都是分类

机器学习模型

$$y = Wx$$

$x$  特征表示向量

$W$  模型参数

$y$  输出向量



# 自然语言处理基本方法



## 文本分类

<https://github.com/bamtercelboo/cnn-lstm-bilstm-deepcnn-clstm-in-pytorch>

## 序列标注

<https://github.com/jiesutd/NCRFpp>

## 结构学习

## 序列到序列

<https://github.com/whr94621/NJUNMT-pytorch>

# 主要内容



- 自然语言处理任务简介
- 文本分类
- 序列标注
- 结构学习
- 序列到序列

# 文本分类



情感分析

主题分类

文本蕴含

阅读理解

关系分类

股票预测

票房分析

.....

# 情感分析



我自己连看了三遍夏洛特烦恼  
正面|负面|中立

# 情感分析



输入:

我自己连看了三遍夏洛特烦恼

# 情感分析



## 传统方法

我自己连看了三遍 夏洛特烦恼

# 情感分析



## 传统方法

我自己连看了三遍夏洛特烦恼

### 1. 特征提取 (T1)

$F = \{T1=\text{我}, T1=\text{自己}, T1=\text{连}, T1=\text{看}, T1=\text{了}, T1=\text{三}, T1=\text{遍}, T1=\text{夏洛特烦恼}\}$

# 情感分析



## 传统方法

我自己连看了三遍夏洛特烦恼

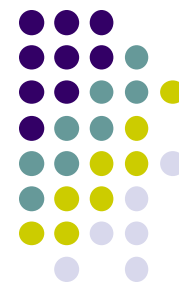
### 1. 特征提取 (T1, T2)

$F = \{T1=我, T1=自己, T1=连, T1=看, T1=了, T1=三, T1=遍, T1=夏洛特烦恼,$

$T2=我自己, T2=自己连, T2=连看, T2=看了, T2=了三, T2=三遍, T2=遍夏洛特烦恼\}$



# 情感分析



## 传统方法

我自己连看了三遍夏洛特烦恼

### 1. 特征提取 (T1, T2)

$F = \{T1=我, T1=自己, T1=连, T1=看, T1=了, T1=三, T1=遍, T1=夏洛特烦恼, T2=我自己, T2=自己连, T2=连看, T2=看了, T2=了三, T2=三遍, T2=遍夏洛特烦恼\}$

### 2. 加载模型 (V, W)

V: 特征词典集合  $x = lookup(F, V) \quad x \in R^{|V|}$  (行向量)

W: 特征极性权重  $W \in R^{3 \times |V|}$

# 情感分析



## 传统方法

我自己连看了三遍 夏洛特烦恼

1. 特征提取 (T1, T2)

2. 加载模型 (V, W)

V: 特征词典集合  $x = \text{lookup}(F, V)$   $x \in R^{|V|}$

W: 特征极性权重  $W \in R^{3 \times |V|}$

3. 预测

$$y = Wx$$

# 情感分析



## 传统方法

我自己连看了三遍夏洛特烦恼

1. 特征提取 (T1, T2)
2. 加载模型 (V, W)
3. 预测  $y = Wx$
4. 参数更新(训练、模型调整)

$$\Delta y = \text{dist}(y, y_g) \quad \Delta W = \Delta y x'$$

$$W = W - \eta \Delta W$$

# 情感分析



我自己连看了三遍夏洛特烦恼

传统方法:

特征提取 (T1, T2)

$F = \{T1=我, T1=自己, T1=连, T1=看, T1=了, T1=三, T1=遍, T1=夏洛特烦恼, T2=我自己, T2=自己连, T2=连看, T2=看了, T2=了三, T2=三遍, T2=遍夏洛特烦恼\}$

特征表:  $\Sigma$  (每个特征赋予一个id)

$\langle 0, \dots, 0, 1, \dots, 0, 0, 1, 1, 0, \dots \rangle$

T1=我

T2=自己连

T2=我自己

稀疏向量表示

- 特征总数有多少, 维度便有多少
- 对于每一个样例, 只有很少几个不为零值

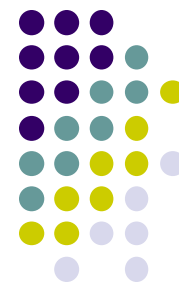
# 深度学习



我自己连看了三遍夏洛特烦恼

$e_1$   $e_2$   $e_3$   $e_4$   $e_5$   $e_6$   $e_7$   $e_8$

# 深度学习



我自己连看了三遍夏洛特烦恼

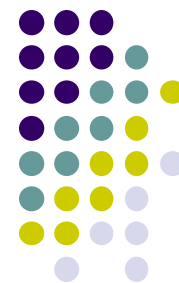
$e_1$     $e_2$     $e_3$     $e_4$     $e_5$     $e_6$     $e_7$     $e_8$

0.2
-1.3
0.5
-10.2
4.3

深度学习的基础：稠密向量表示

- 维数比较低，一般不超过500维
- 基本上都为非零值，实数

# 深度学习



我自己连看了三遍夏洛特烦恼

$e_1$     $e_2$     $e_3$     $e_4$     $e_5$     $e_6$     $e_7$     $e_8$

0.2
-1.3
0.5
-10.2
4.3

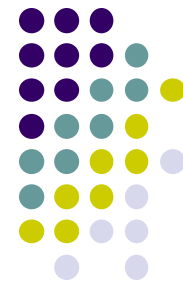
深度学习的基础：稠密向量表示

- 维数比较低，一般不超过500维
- 基本上都为非零值，实数

词向量获取方式

- 随机初始化，本身作为模型参数参与参数学习
- word2vec

# 深度学习

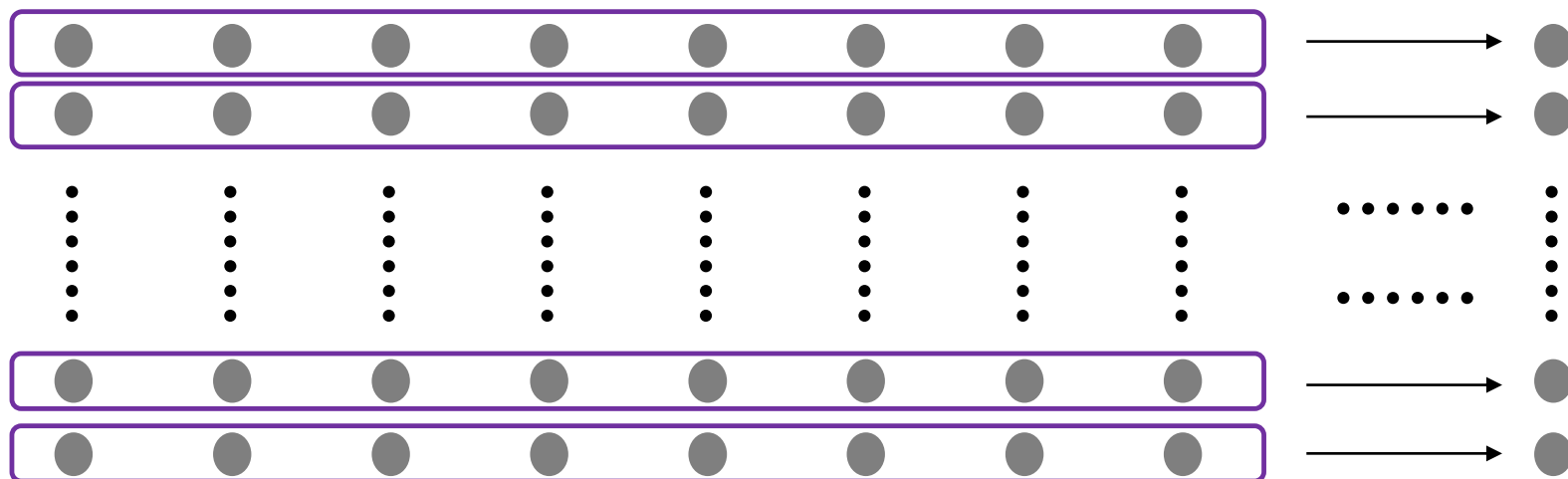


我自己连看了三遍夏洛特烦恼

$e_1$   $e_2$   $e_3$   $e_4$   $e_5$   $e_6$   $e_7$   $e_8$

Pooling:

max  $x$





# 深度学习



我自己连看了三遍夏洛特烦恼

$e_1$   $e_2$   $e_3$   $e_4$   $e_5$   $e_6$   $e_7$   $e_8$

Pooling:  $x = \sum A_i e_i$

预测:  $y = Wx$

# 深度学习



我自己连看了三遍夏洛特烦恼

$e_1$   $e_2$   $e_3$   $e_4$   $e_5$   $e_6$   $e_7$   $e_8$

**Pooling:**  $x = \sum A_i e_i$  max:  $A_{i,j} = \begin{cases} 1, & \text{当 } i = \arg \max_s (h_{s,j}) \\ 0 \end{cases}$

**预测:**  $y = Wx$

# 深度学习



我自己连看了三遍夏洛特烦恼

$e_1$   $e_2$   $e_3$   $e_4$   $e_5$   $e_6$   $e_7$   $e_8$

**Pooling:**  $x = \sum A_i e_i$     min:  $A_{i,j} = \begin{cases} 1, & \text{当 } i = \arg \min_s (h_{s,j}) \\ 0 \end{cases}$

**预测:**  $y = Wx$

# 深度学习



我自己连看了三遍夏洛特烦恼

$e_1$   $e_2$   $e_3$   $e_4$   $e_5$   $e_6$   $e_7$   $e_8$

Pooling:  $x = \sum A_i e_i$       avg:  $A_i = \frac{1}{n}$   
预测:  $y = Wx$

# 深度学习



我自己连看了三遍夏洛特烦恼

$e_1$   $e_2$   $e_3$   $e_4$   $e_5$   $e_6$   $e_7$   $e_8$

Pooling:  $x = \sum A_i e_i$       sum:  $A_i = 1$

预测:  $y = Wx$

# 深度学习



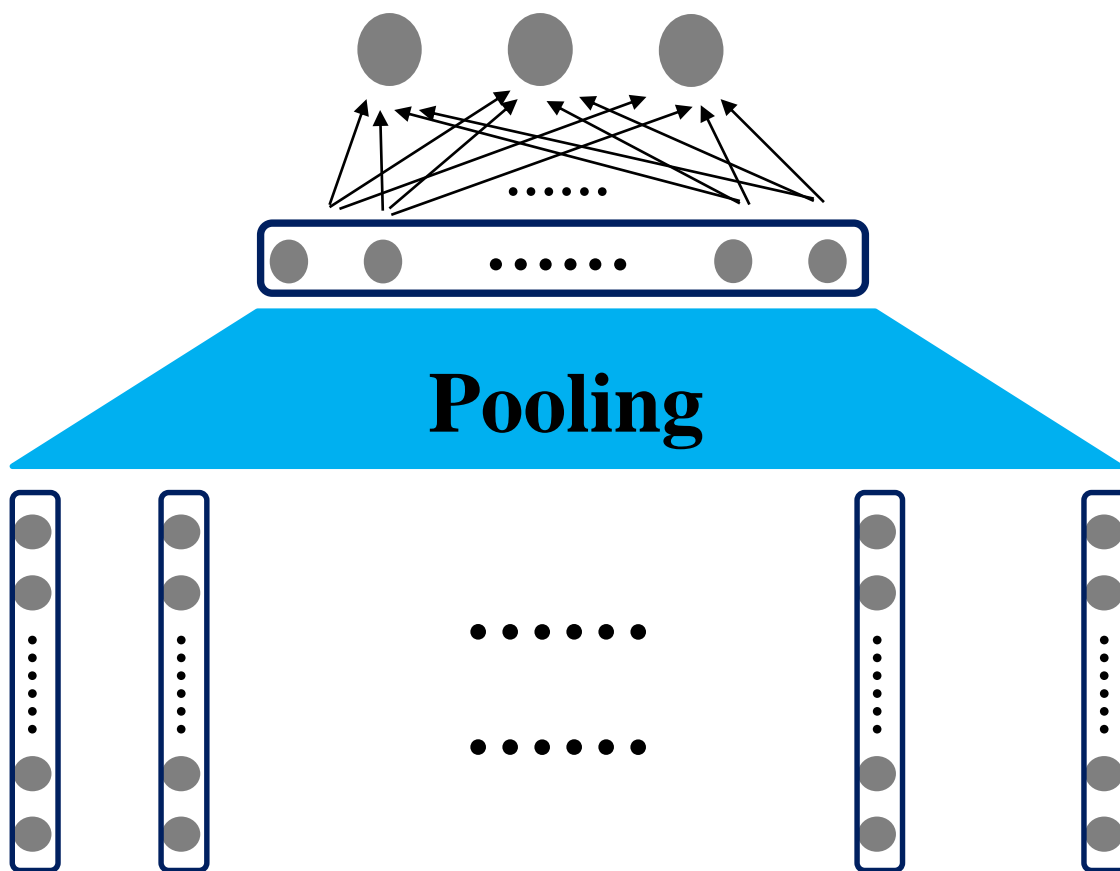
我自己连看了三遍夏洛特烦恼

$e_1$   $e_2$   $e_3$   $e_4$   $e_5$   $e_6$   $e_7$   $e_8$

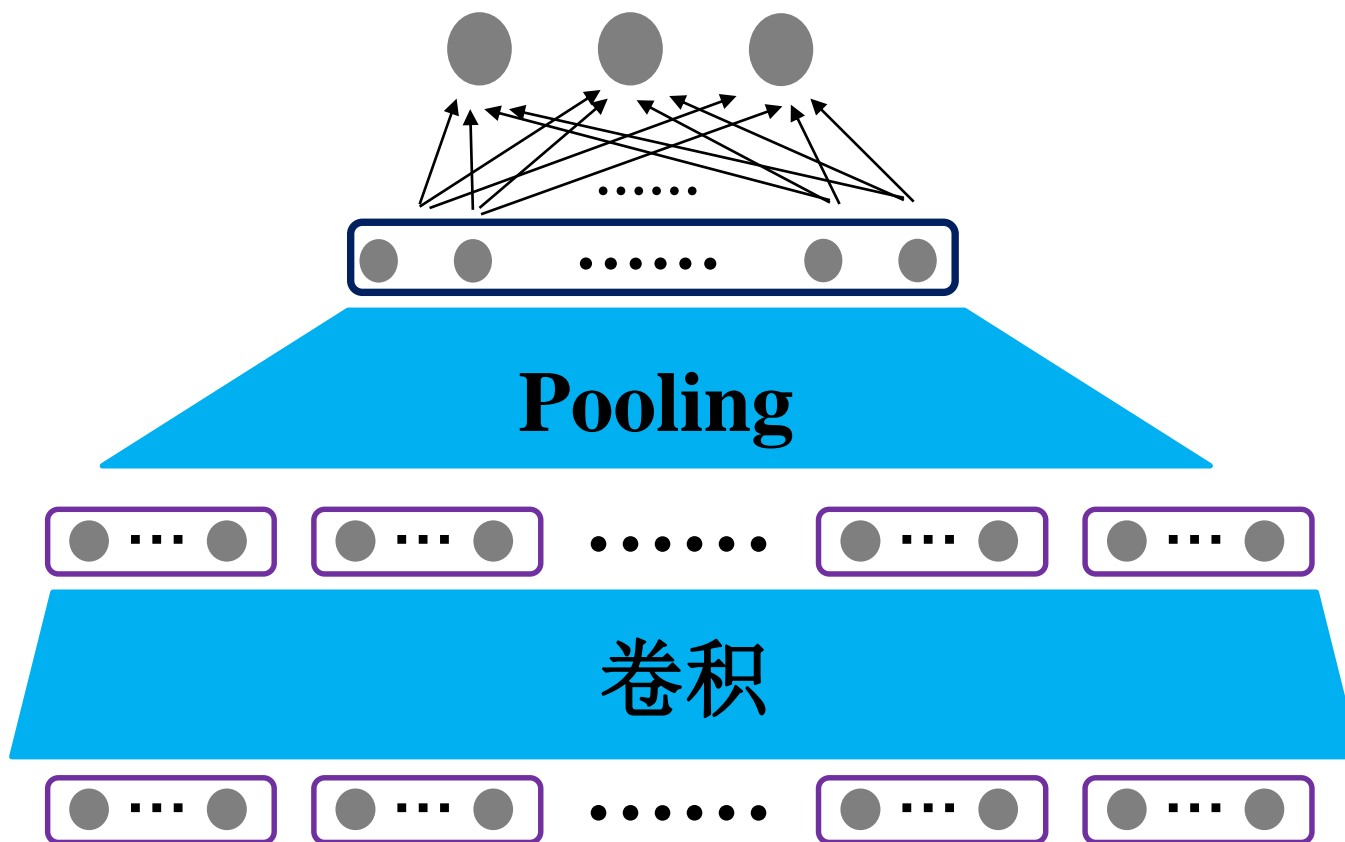
**Pooling:**  $x = \sum A_i e_i$  *gated:*  $A_i \propto \exp(\tanh(W_g e_i + b_g))$

**预测:**  $y = Wx$

# 深度学习



# 深度学习





# 深度学习



我自己连看了三遍夏洛特烦恼

$e_1$   $e_2$   $e_3$   $e_4$   $e_5$   $e_6$   $e_7$   $e_8$

卷积:  $h_i = \tanh(\sum_{j=0}^c U_j e_{i-j})$

Pooling:  $x = \sum A_i h_i$

预测:  $y = Wx$

# 深度学习



卷积:  $h_i = \tanh(\sum_{j=0}^c U_j e_{i-j})$

Pooling:  $x = \sum A_i h_i$

预测:  $y = Wx$

参数更新 (训练): 反向传播

损失:  $\Delta y = \text{dist}(y, y_g)$

# 深度学习



卷积:  $h_i = \tanh(\sum_{j=0}^c U_j e_{i-j})$

Pooling:  $x = \sum A_i h_i$

预测:  $y = Wx$

参数更新 (训练): 反向传播

损失:  $\Delta y = \text{dist}(y, y_g)$

预测:  $\Delta W = \Delta y x' ; \Delta x = W' \Delta y$

$W = W - \eta \Delta W$

$$\Delta x = \Delta y \frac{\partial y}{\partial x}$$

# 深度学习



卷积:  $h_i = \tanh(\sum_{j=0}^c U_j e_{i-j})$

Pooling:  $x = \sum A_i h_i$

预测:  $y = Wx$

参数更新 (训练): 反向传播

损失:  $\Delta y = \text{dist}(y, y_g)$

预测:  $\Delta W = \Delta y x' ; \Delta x = W' \Delta y$

$W = W - \eta \Delta W$

Pooling:  $\Delta h_i = A_i \Delta x'$

$$\Delta x = \Delta y \frac{\partial y}{\partial x}$$

# 深度学习



卷积:  $h_i = \tanh(\sum_{j=0}^c U_j e_{i-j})$

Pooling:  $x = \sum A_i h_i$

预测:  $y = Wx$

参数更新 (训练): 反向传播

损失:  $\Delta y = \text{dist}(y, y_g)$

预测:  $\Delta W = \Delta y x' ; \Delta x = W' \Delta y$

$$W = W - \eta \Delta W$$

Pooling:  $\Delta h_i = A_i \Delta x'$

卷积:  $\Delta U_j = \Delta U_j + \Delta h_i \frac{\partial h_i}{\partial U_j} ; \Delta e_i = \Delta e_i + \Delta h_j \frac{\partial h_j}{\partial e_i}$

$$U = U - \eta \Delta U \quad e = e - \eta \Delta e$$

$$\Delta x = \Delta y \frac{\partial y}{\partial x}$$

# 深度学习



随机初始化所有模型参数  $(-0.01, +0.01)$

for ....

输入一个  $(x, y)$ ,  $y$  转化成为  $p_{y_g}$

$h = \text{NeuralNetwork}(x)$

$y = Wh$

$p_y = \text{softmax}(y)$  概率形式

$\Delta y = \text{dist}(y, y_g) = p_y - p_{y_g}$

反向传播

End  $p_{y_g}$ : + (1, 0, 0); - (0, 0, 1); 中性 (0, 1, 0)

# 深度学习



大数据预训练:

word2vec

我 自 己 连 看 了 三 遍 夏 洛 特 烦 恼

$e_1$     $e_2$     $e_3$     $e_4$     $e_5$     $e_6$     $e_7$     $e_8$

# 深度学习



大数据预训练:

word2vec

我 自 己 连 ?? 了 三 遍 夏 洛 特 烦 恼

$e_1$     $e_2$     $e_3$              $e_5$     $e_6$     $e_7$              $e_8$



# 主要内容



- 自然语言处理任务简介
- 文本分类
- 序列标注
- 结构学习
- 序列到序列

# 序列标注



- 分词
- 词性标注
- 命名实体识别
- 情感要素识别
- 组块句法分析

.....

# 分词



输入： 本次课将介绍自然语言处理相关的基本任务

输出： 本 次 课 将 介 绍 自 然 语 言 处 理 相 关 的 基 本 任 务

# 分词



- 分词歧义

输入： 羽毛球拍卖完了

结果： 羽毛球**球拍** 卖 完 了

输入： 他站起身来

结果： 他 站 **起 身** 来

(他 明天 **起身** 去 北京)

# 分词



输入：羽毛球拍卖完了

结果：羽毛球**球拍** 卖 完 了

分类：  $f(\text{上下文}) \rightarrow$  标签

如何将分词转化成为打标签任务？

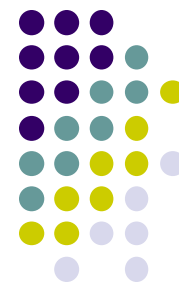
羽 毛 球 拍 卖 完 了

B I I I B B B

两类标签

羽毛球拍 卖 完 了

# 分词



输入： 羽毛球拍卖完了

结果： 羽毛球**球拍** 卖 完 了

分类：  $f(\text{上下文}) \rightarrow$  标签

如何将分词转化成为打标签任务？

羽 毛 **球** 拍 卖 完 了

B I **I** I B B B

两类标签

羽毛球拍 卖 完 了

# 分词



羽 毛 球 拍 卖 完 了

特征提取： 窗口形式

手工模版提取

# 分词



羽 毛 球 拍 卖 完 了

特征提取： 窗口形式

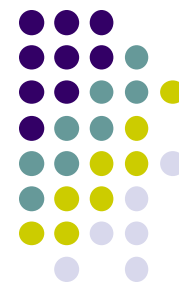
手工模版提取

卷积神经网络：

$$h_i = \tanh(W[e_{i-2} \dots e_{i+2}] + b)$$



# 分词



羽 毛 球 拍 卖 完 了

特征提取： 窗口形式

手工模版提取

卷积神经网络

循环神经网络：  $h_i = \tanh(Wh_{i-1} + b)$

$$h_i = \tanh(Wh_{i+1} + b)$$

# 词性标注



输入：本次 课 将 介绍 自然 语言 处理 相关 的 基本 任务

预测： r n d v n n v v u a n

思考： 与分词的区别在哪里？

# 主要内容



- 自然语言处理任务简介
- 文本分类
- 序列标注
- 结构学习
- 序列到序列

# 结构学习



- 句法分析
- 语义分析
- 信息抽取
- 事件检测
- 意见挖掘

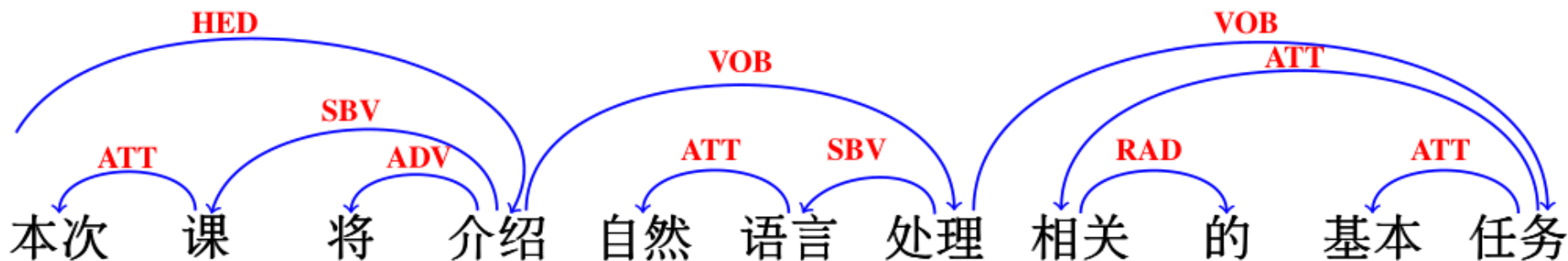
.....

# 结构学习



## ● 句法分析(依存句法分析)

输入：本次 课 将 介绍 自然 语言 处理 相关 的 基本 任务



如何将依存句法分析转换成为分类？

# 依存句法分析



输入：本次课将介绍自然语言处理相关的基本任务

如何将句法分析转换为分类？

基于转移的系统：

**Stack**

|

**Queue**

类别：**action**

# 依存句法分析



输入：本次 课 将 介绍 自然 语言 处理 相关 的 基本 任务

如何将句法分析转换为分类？

本次 课 将 . . .  
← shift

# 依存句法分析



输入：本次 课 将 介绍 自然 语言 处理 相关 的 基本 任务

如何将句法分析转换为分类？

本次		课	将	介绍 · · ·
		← shift		



# 依存句法分析



输入：本次 课 将 介绍 自然 语言 处理 相关 的 基本 任务

如何将句法分析转换为分类？

本次 课

arc-left(ATT)

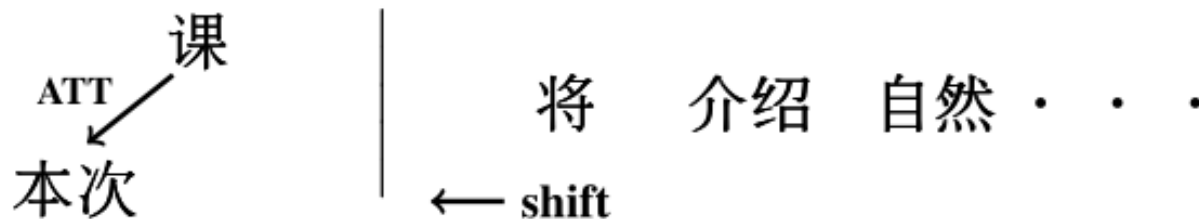
将 介绍 自然 · · ·

# 依存句法分析



输入：本次 课 将 介绍 自然 语言 处理 相关 的 基本 任务

如何将句法分析转换为分类？

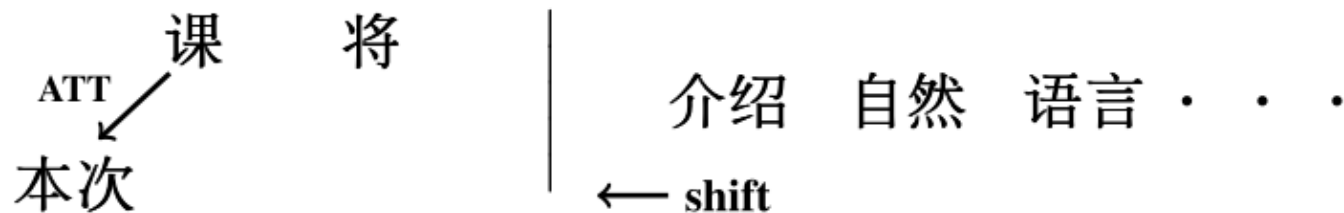


# 依存句法分析



输入：本次 课 将 介绍 自然 语言 处理 相关 的 基本 任务

如何将句法分析转换为分类？

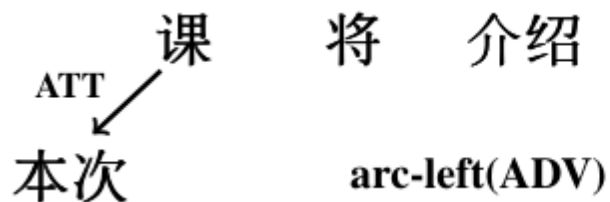


# 依存句法分析



输入：本次 课 将 介绍 自然 语言 处理 相关 的 基本 任务

如何将句法分析转换为分类？



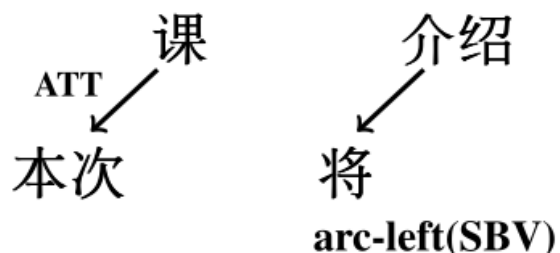
自然 语言 处理 . . .

# 依存句法分析



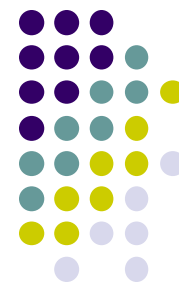
输入：本次 课 将 介绍 自然 语言 处理 相关 的 基本 任务

如何将句法分析转换为分类？



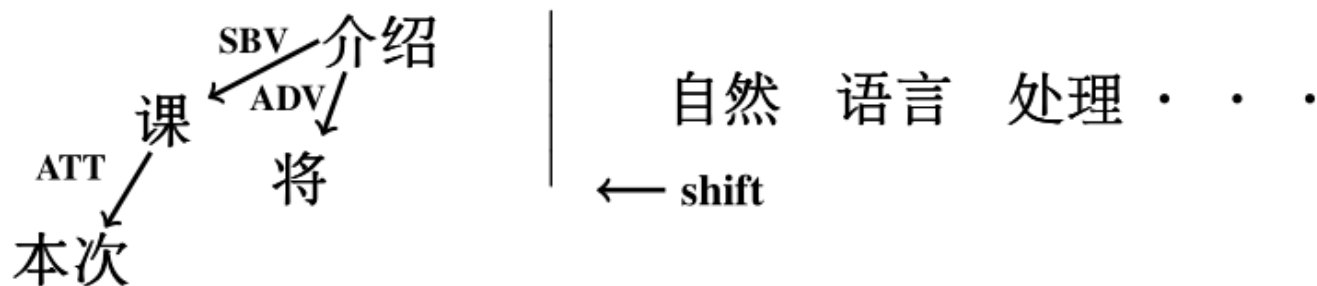
自然 语言 处理 . . .

# 依存句法分析



输入：本次 课 将 介绍 自然 语言 处理 相关 的 基本 任务

如何将句法分析转换为分类？

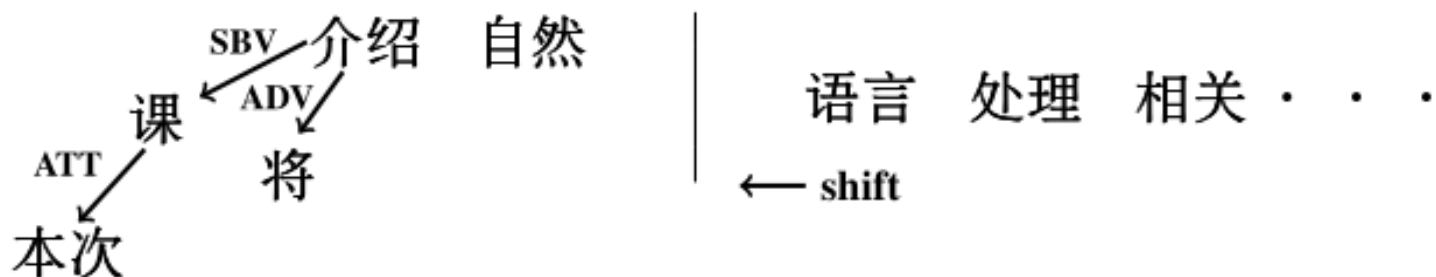


# 依存句法分析

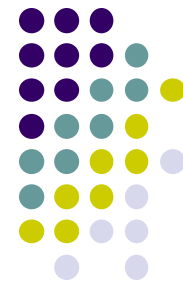


输入：本次 课 将 介绍 自然 语言 处理 相关 的 基本 任务

如何将句法分析转换为分类？

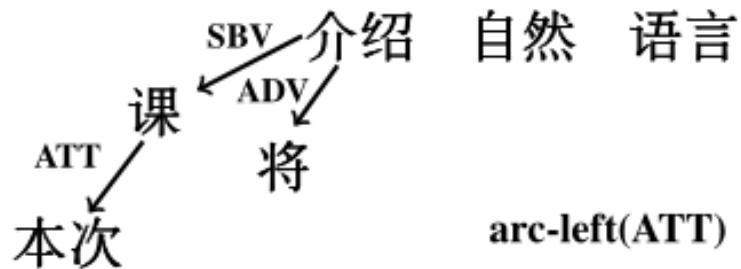


# 依存句法分析



输入：本次 课 将 介绍 自然 语言 处理 相关 的 基本 任务

如何将句法分析转换为分类？



处理 相关 的 . . .

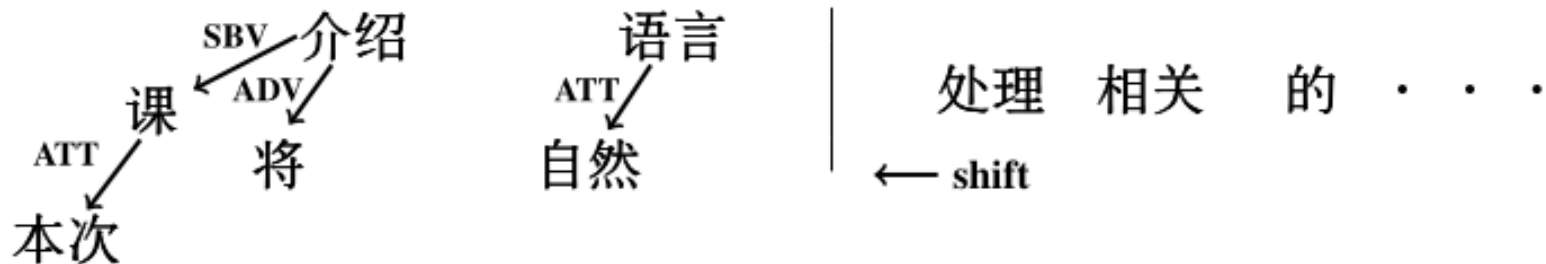


# 依存句法分析



输入：本次 课 将 介绍 自然 语言 处理 相关 的 基本 任务

如何将句法分析转换为分类？

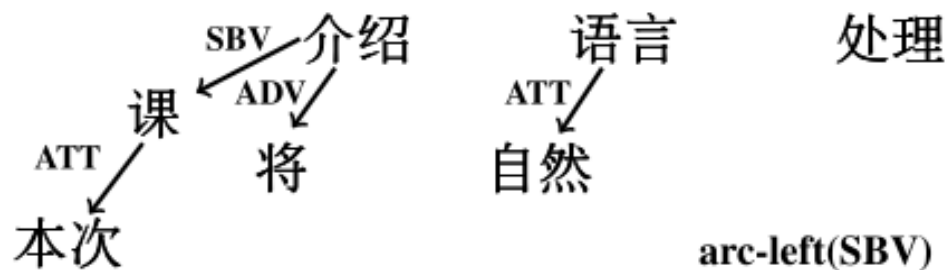


# 依存句法分析



输入：本次 课 将 介绍 自然 语言 处理 相关 的 基本 任务

如何将句法分析转换为分类？



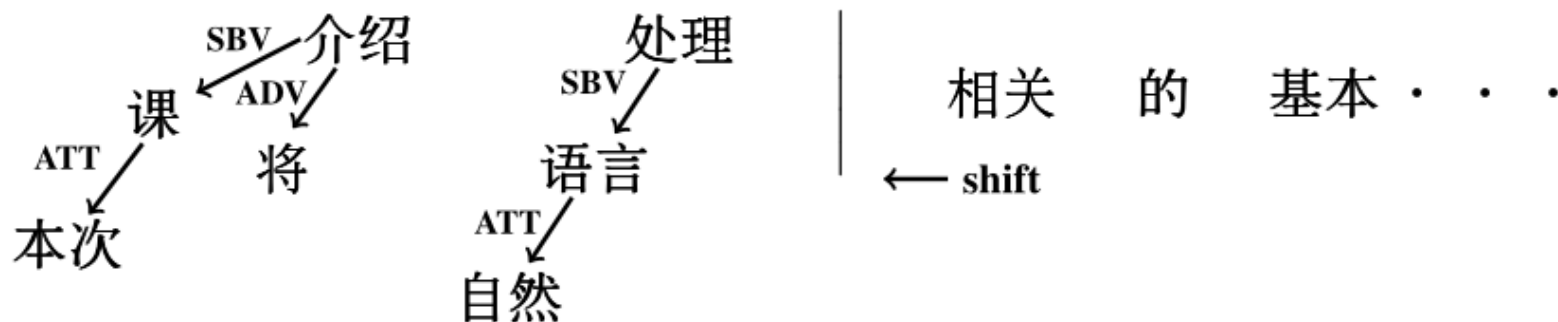
相关 的 基本 . . .

# 依存句法分析



输入：本次 课 将 介绍 自然 语言 处理 相关 的 基本 任务

如何将句法分析转换为分类？

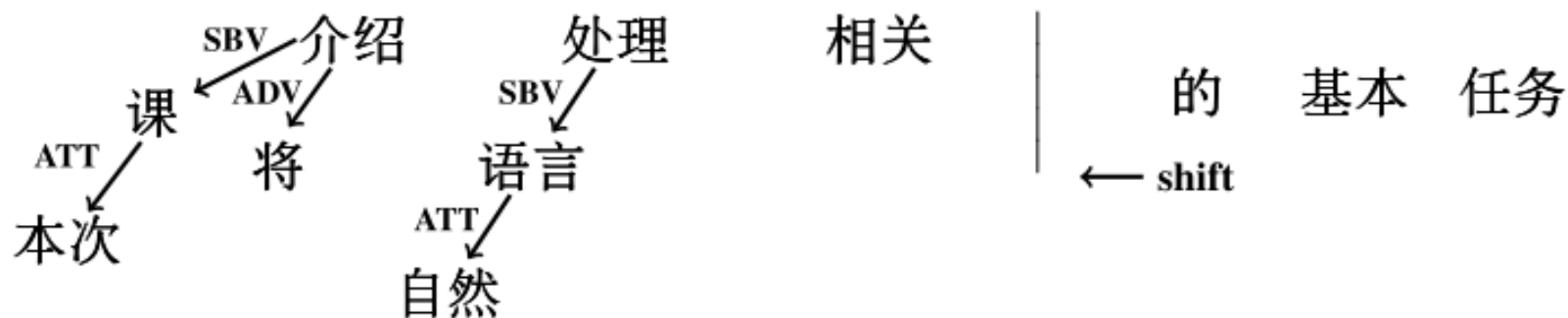


# 依存句法分析



输入：本次 课 将 介绍 自然 语言 处理 相关 的 基本 任务

如何将句法分析转换为分类？

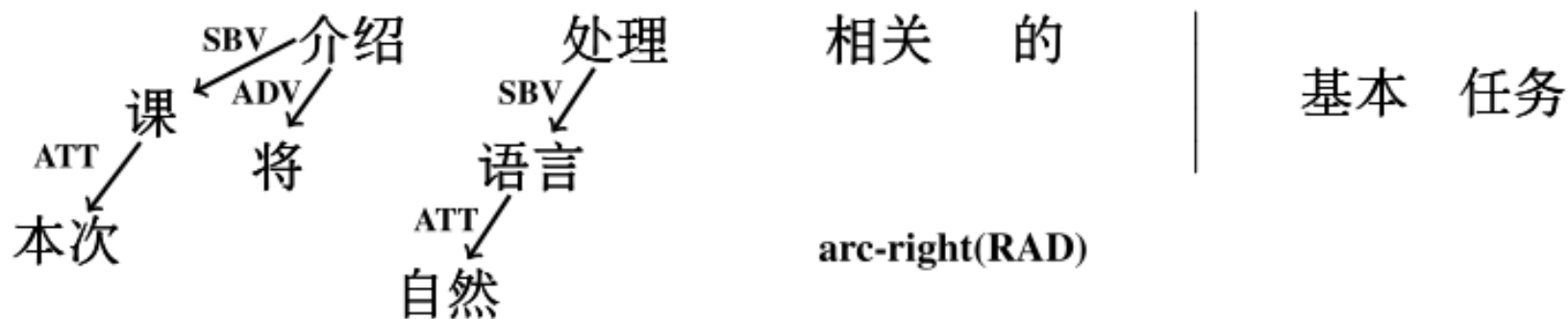


# 依存句法分析



输入：本次 课 将 介绍 自然 语言 处理 相关 的 基本 任务

如何将句法分析转换为分类？

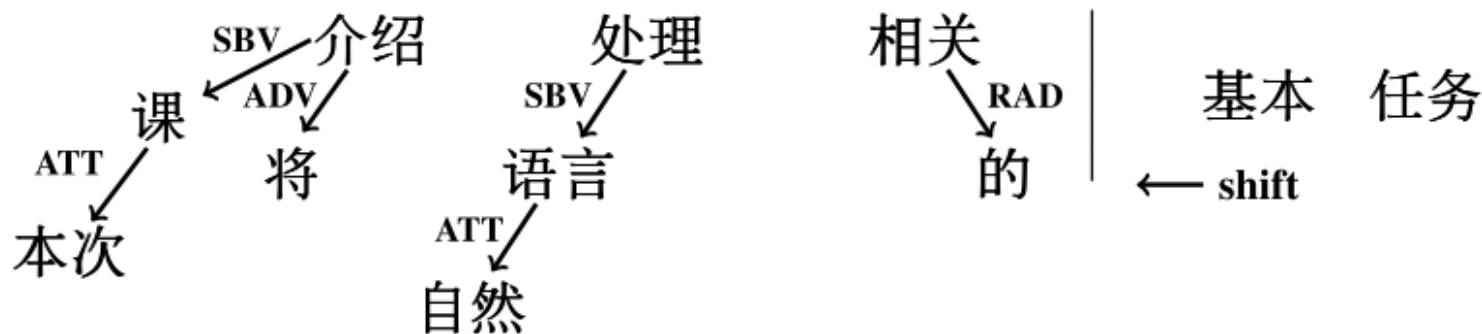


# 依存句法分析



输入：本次 课 将 介绍 自然 语言 处理 相关 的 基本 任务

如何将句法分析转换为分类？

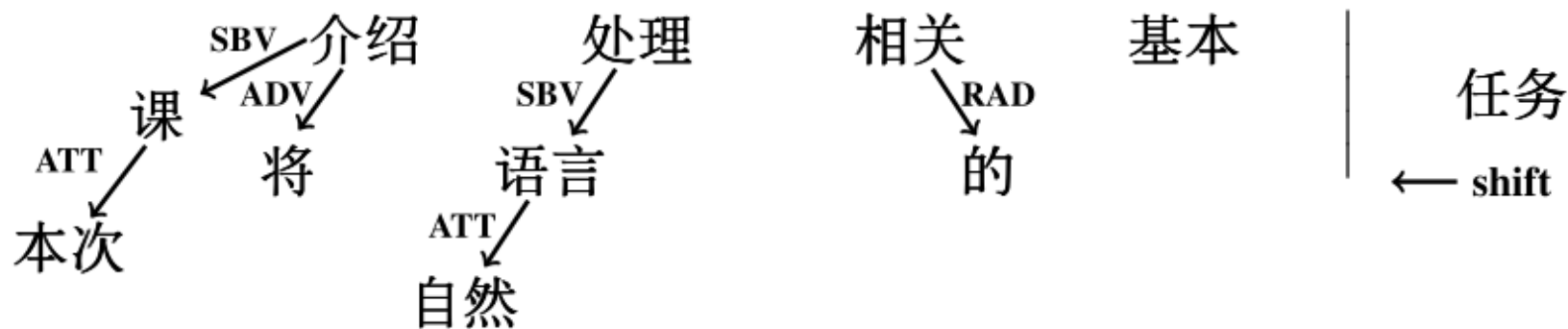


# 依存句法分析



输入：本次 课 将 介绍 自然 语言 处理 相关 的 基本 任务

如何将句法分析转换为分类？

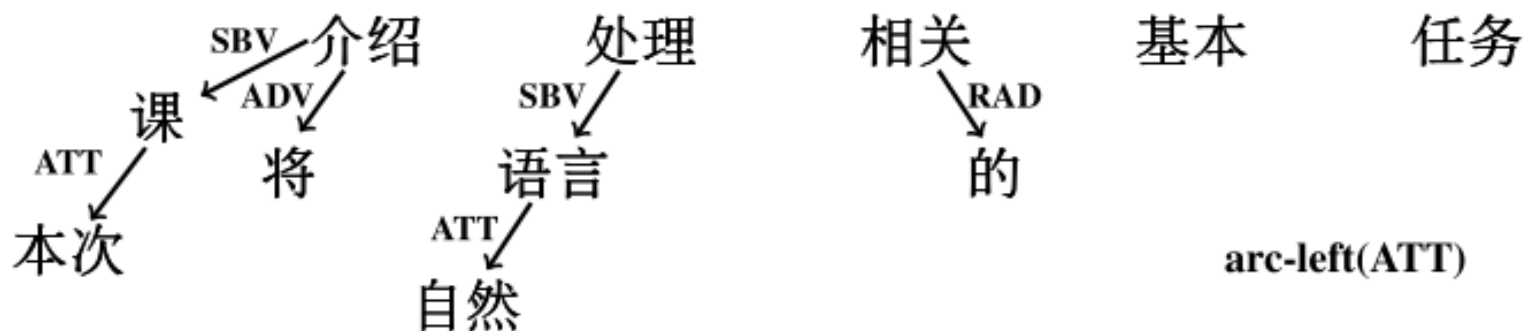




# 依存句法分析

输入：本次 课 将 介绍 自然 语言 处理 相关 的 基本 任务

如何将句法分析转换为分类？





# 依存句法分析



输入：本次 课 将 介绍 自然 语言 处理 相关 的 基本 任务

如何将句法分析转换为分类？

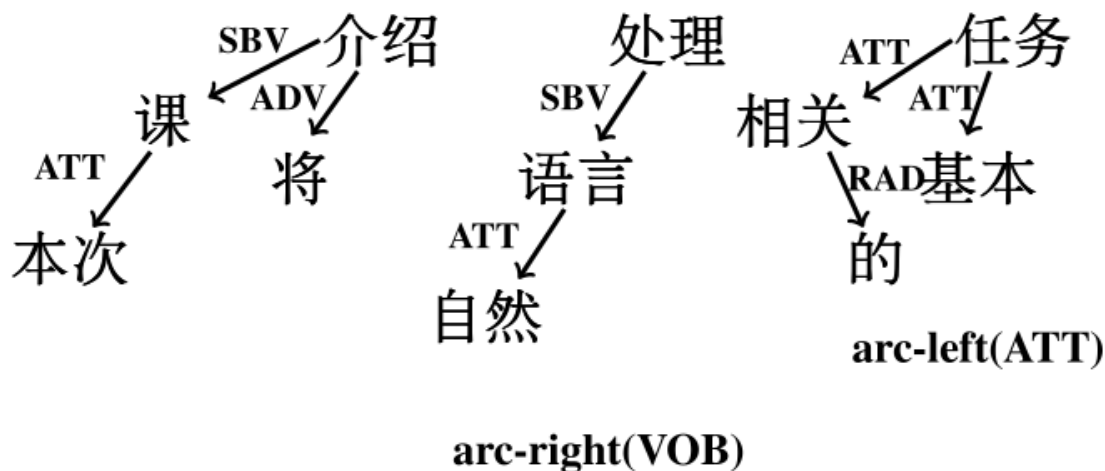


# 依存句法分析



输入：本次 课 将 介绍 自然 语言 处理 相关 的 基本 任务

如何将句法分析转换为分类？

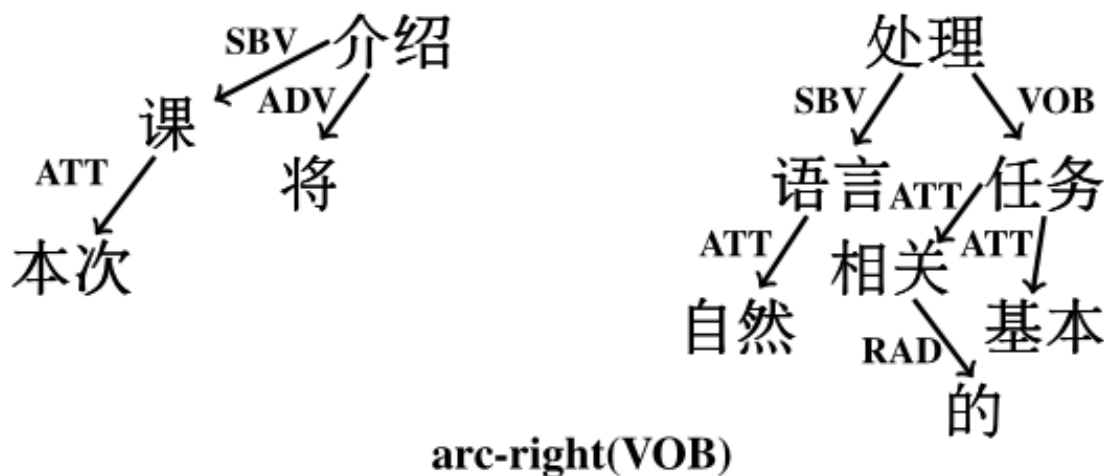




# 依存句法分析

输入：本次 课 将 介绍 自然 语言 处理 相关 的 基本 任务

如何将句法分析转换为分类？

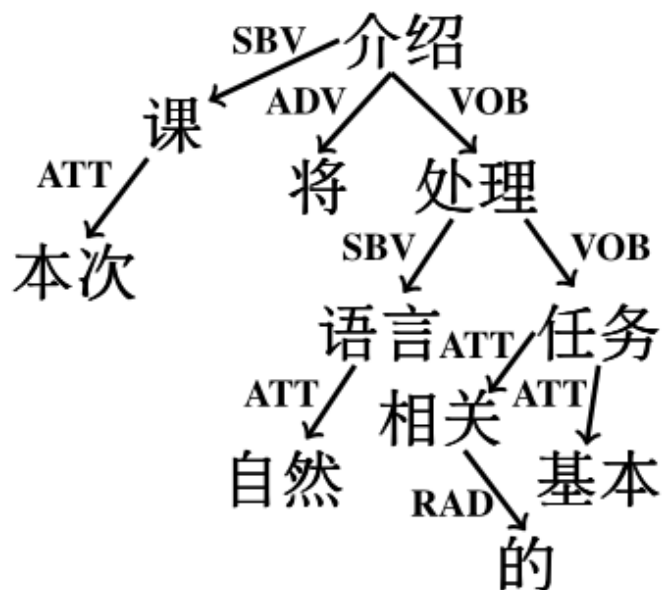




# 依存句法分析

输入：本次课将介绍自然语言处理相关的基本任务

如何将句法分析转换为分类？



# 依存句法分析



输入：本次课将介绍自然语言处理相关的基本任务

如何将句法分析转换为分类？

每一步都是分类：

从所有可能的动作中选取一个目标动作

- **arc-left(\*)**
- **arc-right(\*)**
- **shift**

# 依存句法分析



输入：本次课将介绍自然语言处理相关的基本任务

如何将句法分析转换为分类？

每一步都是分类：

- **arc-left(\*)**
- **arc-right(\*)**
- **shift**

$f(\text{上下文特征}) \rightarrow \text{标签}$

如何提取特征？

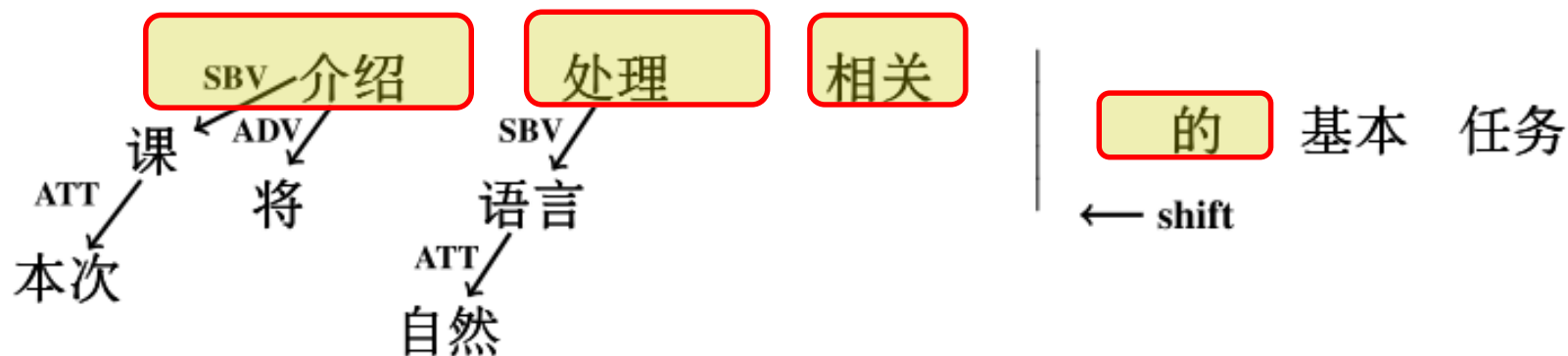
# 依存句法分析



输入：本次 课 将 介绍 自然 语言 处理 相关 的 基本 任务

如何将句法分析转换为分类？

如何提取特征？



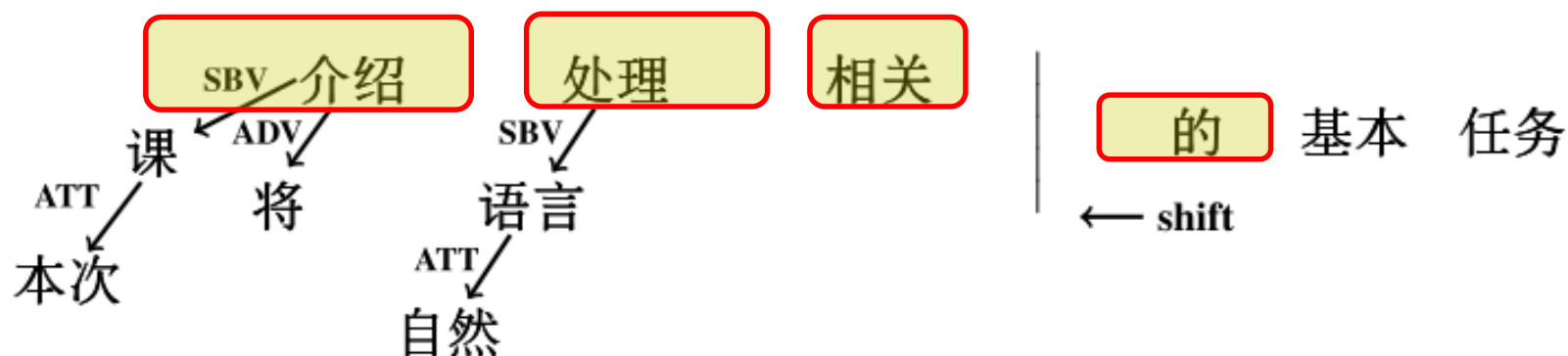
# 依存句法分析



输入：本次 课 将 介绍 自然 语言 处理 相关 的 基本 任务

如何将句法分析转换为分类？

如何提取特征？



人工模板、神经网络（循环神经网络）



# 主要内容



- 自然语言处理任务简介
- 文本分类
- 序列标注
- 结构学习
- 序列到序列

# 序列到序列



机器翻译

对话聊天

客服系统

文本摘要

问答

.....

# 机器翻译



输入: 马拉多纳 则 斥责 老 对手 英格兰 在 裁判 判罚 中 获利

输出: Maradona rebuked his old rival , England , for making profits in the referee 's decision .

# 机器翻译



输入: 马拉多纳 则 斥责 老 对手 英格兰 在 裁判 判罚 中 获利

输出: Maradona rebuked his old rival , England , for making profits in the referee 's decision .

如何分类?

# 机器翻译



输入: 马拉多纳 则 斥责 老 对手 英格兰 在 裁判 判罚 中 获利

输出: **Maradona**

如何分类?

# 机器翻译



输入: 马拉多纳 则 斥责 老 对手 英格兰 在 裁判 判罚 中 获利

输出: Maradona rebuked

如何分类?

# 机器翻译



输入: 马拉多纳 则 斥责 老 对手 英格兰 在 裁判 判罚 中 获利

输出: Maradona rebuked his

如何分类?

# 机器翻译



输入: 马拉多纳 则 斥责 老 对手 英格兰 在 裁判 判罚 中 获利

输出: Maradona rebuked his old

如何分类?



# 机器翻译



输入: 马拉多纳 则 斥责 老 对手 英格兰 在 裁判 判罚 中 获利



如何分类?

# 机器翻译



输入: 马拉多纳 则 斥责 老 对手 英格兰 在 裁判 判罚 中 获利

输出: Maradona rebuked his old rival , England , for making profits in the referee 's decision .

如何分类?

# 机器翻译



输入: 马拉多纳 则 斥责 老 对手 英格兰 在 裁判 判罚 中 获利

输出: Maradona rebuked his old rival , England , for making profits in the referee 's decision . **<eos>**

如何分类?

# 机器翻译



输入: 马拉多纳 则 斥责 老 对手 英格兰 在 裁判 判罚 中 获利

输出: Maradona rebuked his **old**

特征提取?

# 机器翻译

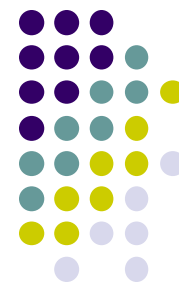


输入: 马拉多纳 则 斥责 老 对手 英格兰 在 裁判 判罚 中 获利  
 $h_1$   $h_2$   $h_3$   $h_4$   $h_5$   $h_6$   $h_7$   $h_8$   $h_9$   $h_{10}$   $h_{11}$

输出: Maradona rebuked his old

特征提取? 编码(Encoder): 循环神经网络

# 机器翻译



输入: 马拉多纳 则 斥责 老 对手 英格兰 在 裁判 判罚 中 获利

$h_1$   $h_2$   $h_3$   $h_4$   $h_5$   $h_6$   $h_7$   $h_8$   $h_9$   $h_{10}$   $h_{11}$

$a_1$   $a_2$   $a_3$   $a_4=1$   $a_5$   $a_6$   $a_7$   $a_8$   $a_9$   $a_{10}$   $a_{11}$

输出: Maradona rebuked his old

特征提取? 编码(Encoder): 循环神经网络

解码(Decoder): 注意力机制+ 分类

$$s_i = V_i \tanh(Wh_i + Uc_i + b)$$
$$a_i \propto \exp(s_i)$$

# 机器翻译



输入: 马拉多纳 则 斥责 老 对手 英格兰 在 裁判 判罚 中 获利  
 $h_1$   $h_2$   $h_3$   $h_4$   $h_5$   $h_6$   $h_7$   $h_8$   $h_9$   $h_{10}$   $h_{11}$   
 $a_1$   $a_2$   $a_3$   $a_4=1$   $a_5$   $a_6$   $a_7$   $a_8$   $a_9$   $a_{10}$   $a_{11}$

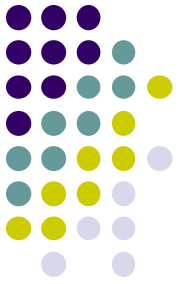
输出: Maradona rebuked his old

单向循环神经网络

特征提取? 编码(Encoder): 循环神经网络

解码(Decoder): 注意力机制 + 分类

$$s_i = V_i \tanh(Wh_i + Uc_i + b)$$
$$a_i \propto \exp(s_i)$$



**The End !**  
**Q/A??**