



中国科学院大学
University of Chinese Academy of Sciences

自然语言处理 **Natural language Processing**

授课教师：胡玥

2024.12



课程编码： 180086081203P2002H 课程名称： 自然语言处理 授课团队 黄河燕、胡玥 、张仰森



第 9 章 NLP基础任务

分类+匹配

1/3

内 容 提 要

9.0 概述

9.1 文本分类

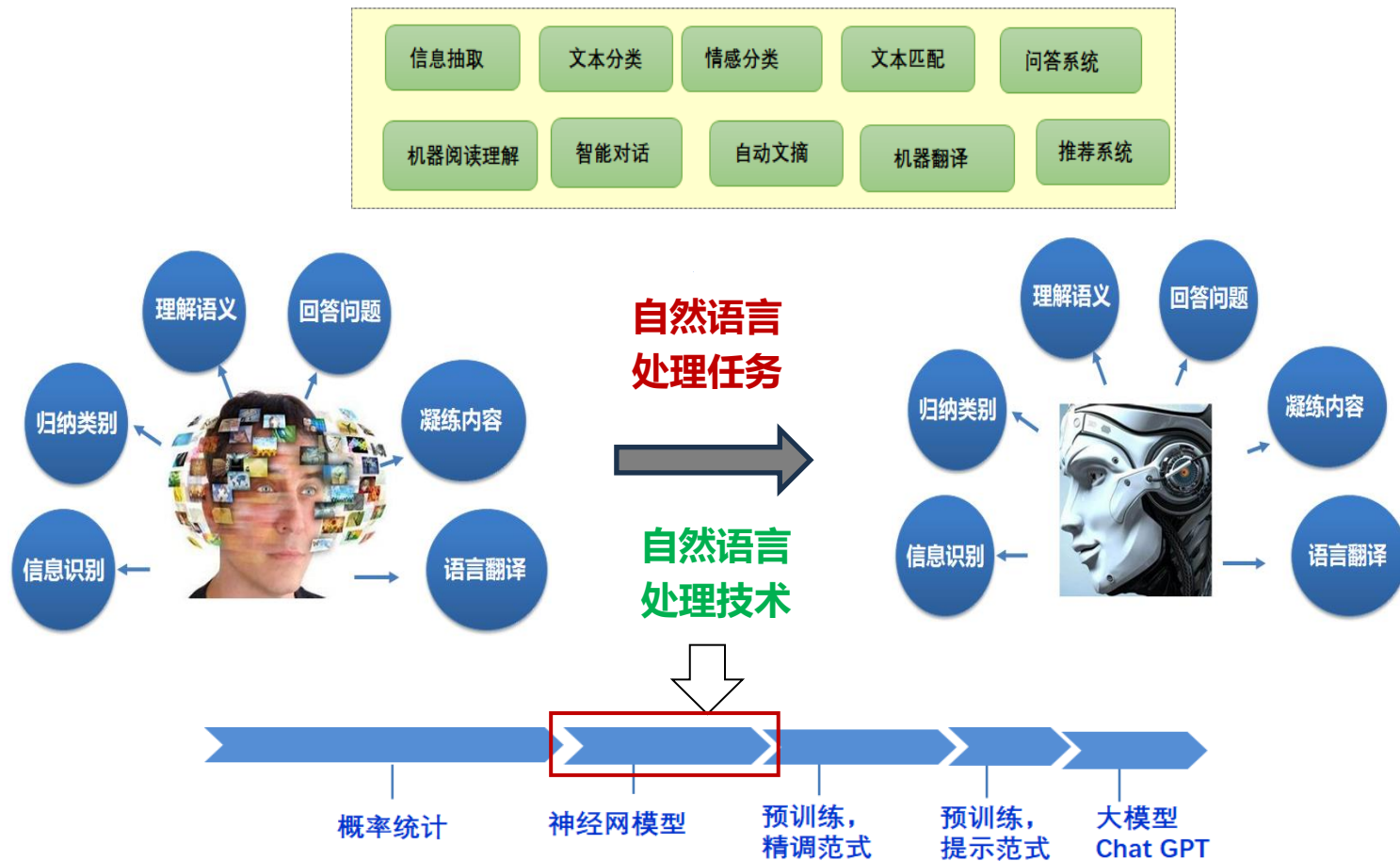
9.2 文本匹配

9.3 序列标注

9.4 序列生成

9.0 概述

智能化语言处理系统



9.0 概述

信息抽取

文本分类

情感分类

文本匹配

问答系统

机器阅读理解

智能对话

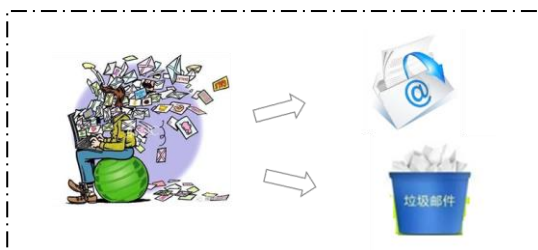
自动文摘

机器翻译

推荐系统

■ 文本分类任务

利用计算机对大量的文档按照分类标准实现自动归类



分类任务	输入	任务建模	输出
	文本序列	分类模型	类别标签 类别标签根据任务定

9.0 概述

信息抽取

文本分类

情感分类

文本匹配

问答系统

机器阅读理解

智能对话

自动文摘

机器翻译

推荐系统

■ 信息抽取任务

信息抽取: 从指定文档中或者海量文本中提取出用户感兴趣的信息。

- 实体识别与抽取
- 实体消歧
- 关系抽取
- 事件抽取

9.0 概述

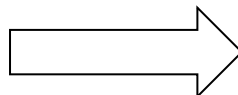
● 实体识别与抽取 （在给定的序列中找到目标序列）

非结构化文本

2011年7月25日，在上海举办的游泳世锦赛上，年仅15岁的叶诗文的以2分08秒90的成绩勇夺女子200米混合泳冠军，成为最年轻的单项世界冠军获得者。

实体抽取

时间/地名
/人名/...



2011年7月25日，在上海举办的游泳世锦赛上，年仅15岁的叶诗文的以2分08秒90的成绩勇夺女子200米混合泳冠军，成为最年轻的单项世界冠军获得者。

标注任务	输入	任务建模	输出
	非结构化文本序列	序列标注模型	目标片段 一般用序列标注方法

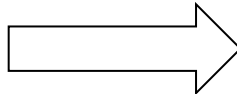
9.0 概述

● 实体间关系抽取

关系抽取：从文本中识别抽取实体及实体之间的关系。抽取结果一般以三元组形式呈现

比尔·盖茨作为微软公司的创始人，以他的创新思维和领导力改变了整个科技行业。他在推动科技进步方面取得了巨大成就

实体间
关系抽取



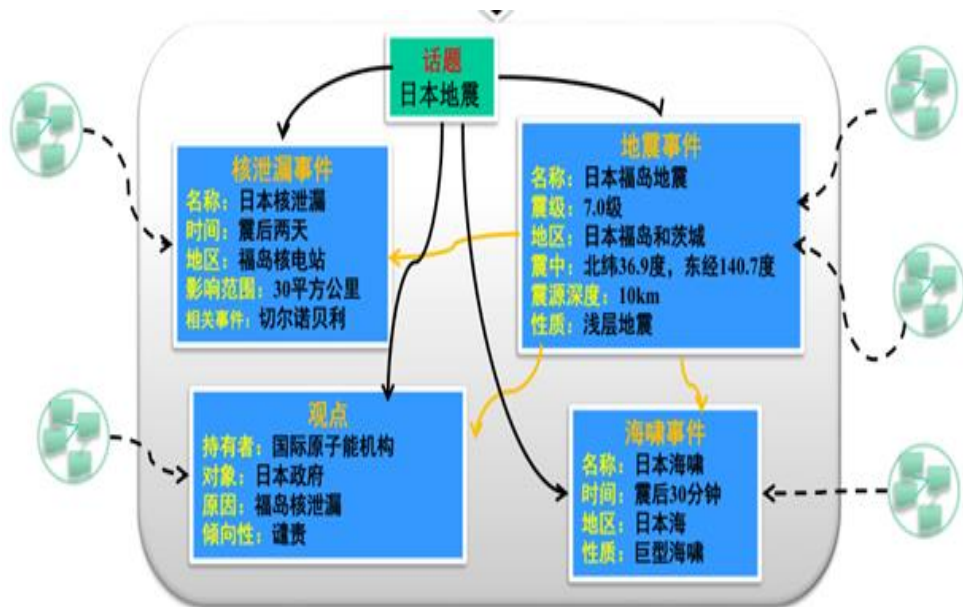
1. 实体：比尔盖茨, 微软
2. 比尔盖茨和微软的关系：
CEO(比尔盖茨, 微软)

关系抽取任务	输入	任务建模	输出
	非结构化文本序列	流水线： 序列标注+分类	1.目标实体(片段) 2. 及之间关系
		端到端： 统一建模	

9.0 概述

● 事件抽取任务

2011年4月11日17点16分，日本东北部的福岛和茨城地区发生里氏7.0级强烈地震(震中北纬36.9度东经140.7度，即福岛西南30公里左右的地方、震源深度10公里，属于浅层地震)当局已经发布海啸预警震后约30分钟后在日本海地区发生巨型海啸，同时造成福岛核电站出现核泄漏，震后第十天，国际原子能机构对于日本政府反应迟钝进行了谴责



9.0 概述

● 事件抽取任务

输入：毛泽东1893年出生于湖南湘潭

1. 触发词检测：（标注+分类）

触发词：出生

触发词类别：life / Be-Born

事件模板

Type/Subtype	Template
Life/Be-Born	Person, Time-Within, Place
Business/Merge-Org	Org, Time, Place
Contact/Meet	Entity, Time, Duration, Place
<i>Transport:</i>	Origin, Artifact, Transporter
-- --	Vehicle, Destination, Agent,
....

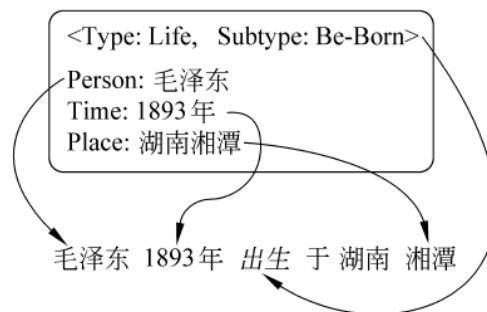
2. life / Be-Born 类模板：

Person:

Time:

Place:

3. 按模板抽取角色：（标注）



输出：事件：

Trigger: 出生

Type: Life, Subtype: Be-Born

Person: 毛泽东

Time: 1893年

Place: 湖南湘潭

任务涉及模型：标注+分类

9.0 概述

信息抽取

文本分类

情感分类

文本匹配

问答系统

机器阅读理解

智能对话

自动文摘

机器翻译

推荐系统

■ 自动文摘

利用计算机自动地从原始文献中提取全面准确地反映文献中心内容地简单连贯的短文。

原文

我非常期待去中国，这是我的一个梦想。我想了解中国美食、文化，想看看这些与国外的中国文化、美食有什么相同和不同的地方。我还想去亲身体验一下中国高山滑雪的场地。在我的职业生涯中，我从未这样期待去中国北京参赛，我对我们墨西哥最终能去中国参赛感到非常开心

摘要

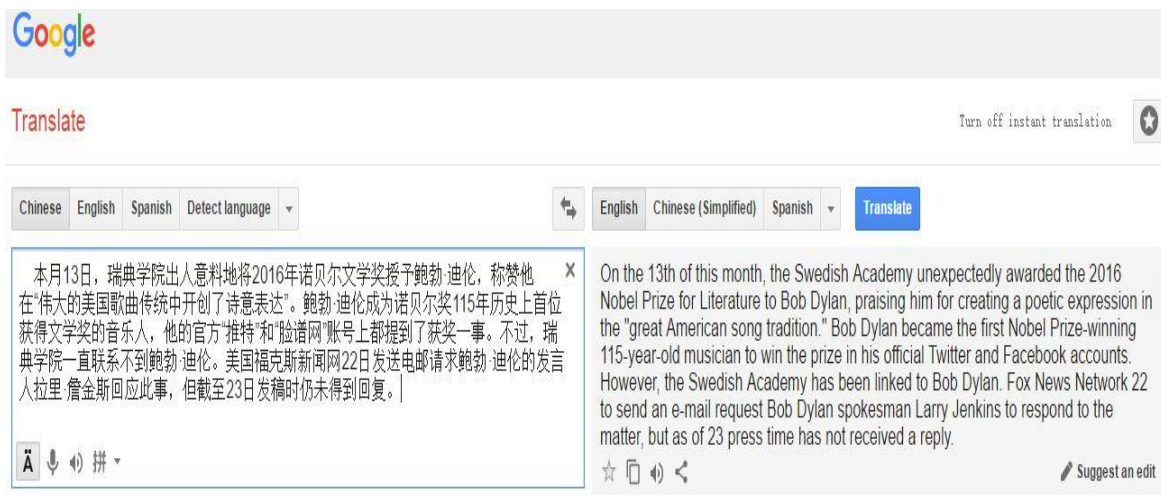
我期待去中国，了解中国美食、文化，并且想去亲身体验一下中国高山滑雪的场地。在我的职业生涯中，能去中国参赛感到非常开心

文摘任务	输入	任务建模	输出
	原文(文本序列)	生成模型	摘要(文本序列)

9.0 概述

■ 机器翻译 (machine translation, MT) 任务

利用计算机把一种语言(源语言, source language) 翻译成另一种语言(目标语言, target language)的技术

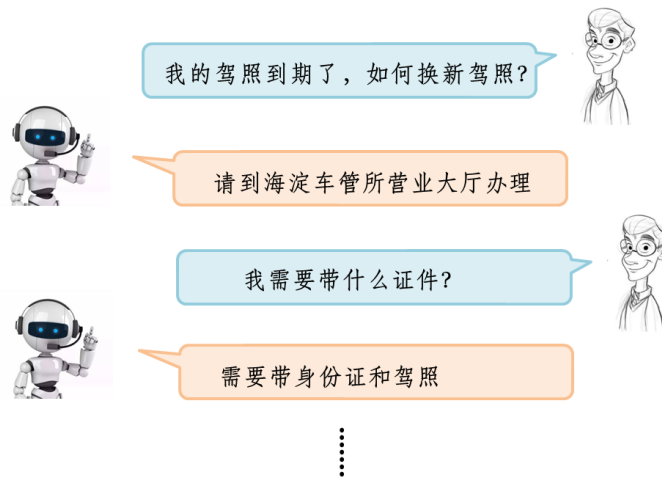


翻译任务	输入	任务建模	输出
	源语(文本序列)	生成模型	目标语(文本序列)

9.0 概述

■ 问答(QA)/对话系统

通过计算机系统对人提出的问题的理解，利用自动推理等手段，在有关知识资源中自动求解答案并做出相应的回答



问答任务	输入	任务建模	输出
	问题(文本序列)	生成模型	回复(文本序列)

9.0 概述

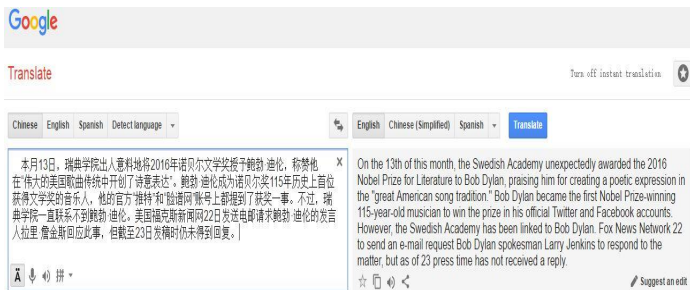
■ 自动文摘/机器翻译/问答(QA)/对话系统 综合

原文

我非常期待去中国，这是我的一个梦想。我了解中国美食、文化，想看看这些与国外的中国文化、美食有什么相同和不同的地方。我还想去亲身体验一下中国高山滑雪的场地。在我的职业生涯中，我从未这样期待去中国北京参赛，我对我们墨西哥最终能去中国参赛感到非常开心

摘要

我期待去中国，了解中国美食、文化，并且想去亲身体验一下中国高山滑雪的场地。在我的职业生涯中，能去中国参赛感到非常开心



我的驾照到期了，如何换新驾照？



请到海淀车管所营业大厅办理



我需要带什么证件？



需要带身份证和驾照

⋮

生成任务	输入	任务建模	输出
	文本序列	生成模型	文本序列

9.0 概述

信息抽取

文本分类

情感分类

文本匹配

问答系统

机器阅读理解

智能对话

自动文摘

机器翻译

推荐系统

■ 文本匹配

广义的将研究两段文本间关系的问题定义为“**文本匹配**”问题，**匹配**含义根据任务的不同有不同的定义。很多自然语言处理的任务都会涉及文本匹配问题

- 复述识别 (paraphrase identification)

两个句子“**感冒了是否要吃药**”和“**感冒了要吃什么药**”

问：两个句子是否表达同样的意思？

- 文本蕴含识别 (Textual Entailment)

两个句子“**我正在上海旅游**”和“**我正在八达岭长城**”

问：这两句话是什么关系？

匹配任务	输入	任务建模	输出
	二段文本序列	匹配模型	二者关系 一般类别标签

9.0 概述

信息抽取

文本分类

情感分类

文本匹配

问答系统

机器阅读理解

智能对话

自动文摘

机器翻译

推荐系统

■ 机器阅读理解

段落

工商协进会报告，12月消费者信心上升到78.1，明显高于11月的72。另据《华尔街日报》报道，2013年是1995年以来美国股市表现最好的一年。这一年里，投资美国股市的明智做法是追着“傻钱”跑。所谓的“傻钱”策略，其实就是买入并持有美国股票这样的普通组合。这个策略要比对冲基金和其它专业投资者使用的更为复杂的投资方法效果好得多。

问题1：什么是傻钱策略？

答案：买入并持有美国股票这样的普通组合

问题2：12月的消费者信心指数是多少？

答案：78.1

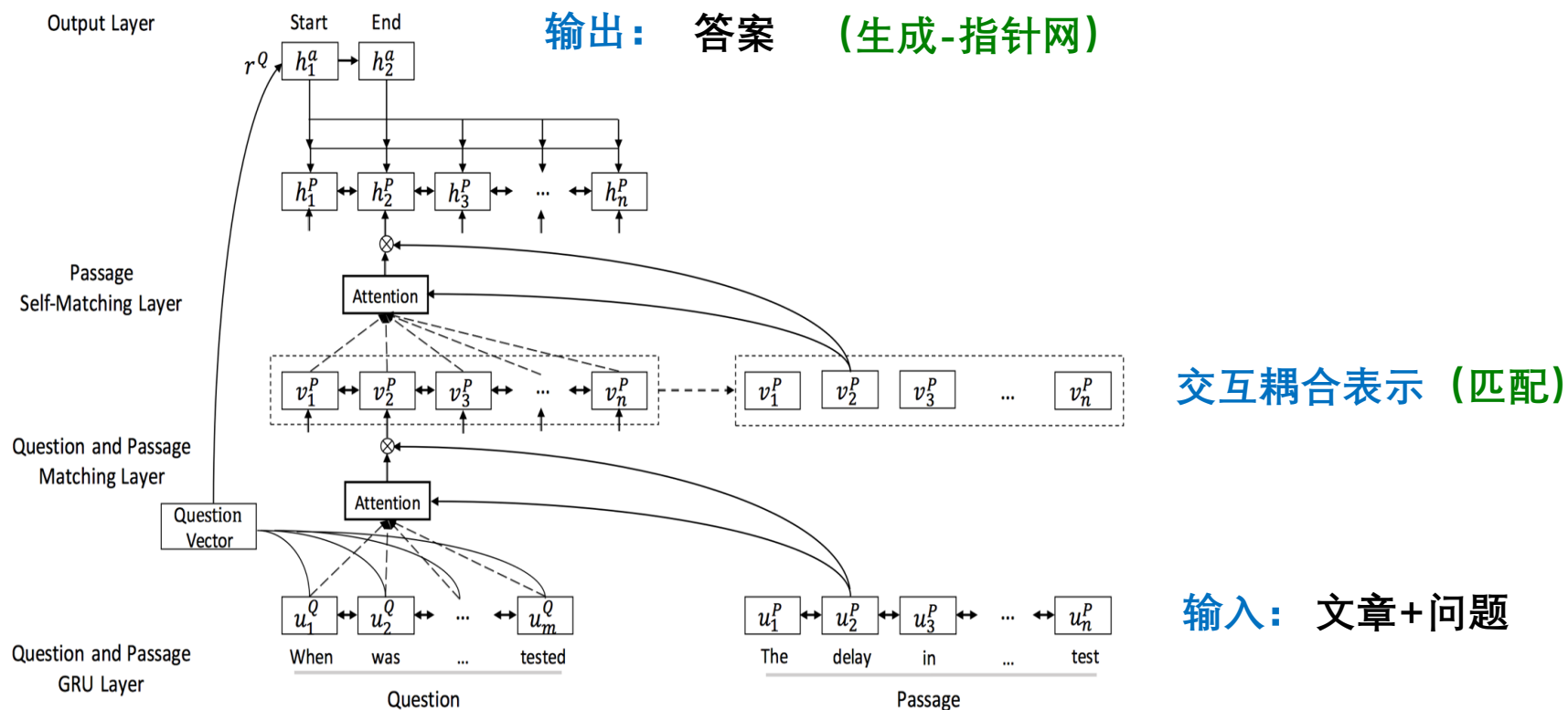
问题3：消费者信心指数由什么机构发布？

答案：工商协进会

阅 读 理 解	输入	任务建模	输出
	二段文本序列 (文章+问题)	匹配模型+ 生成模型	答案 (文章中的片段)

9.0 概述

例：R-NET 模型



9.0 概述

■ 基本任务类型归纳

分类任务	输入	任务建模	输出
	文本序列	分类模型	类别标签 类别标签根据任务定

标注任务	输入	任务建模	输出
	非结构化文本序列	序列标注模型	目标片段 一般用序列标注方法

生成任务	输入	任务建模	输出
	文本序列	生成模型	文本序列

匹配任务	输入	任务建模	输出
	二段文本序列	匹配模型	二者关系 一般类别标签

2, 3, 4 范式



生成模型

分类任务	输入	任务建模	输出
	文本序列	分类模型	类别标签 类别标签根据任务定
标注任务	输入	任务建模	输出
	非结构化文本序列	序列标注模型	目标片段 一般用序列标注方法
生成任务	输入	任务建模	输出
	文本序列	生成模型	文本序列
匹配任务	输入	任务建模	输出
	二段文本序列	匹配模型	二者关系 一般类别标签

5 范式

问题：哪类模型可以完成所有的任务？

内 容 提 要

9.0 概述

9.1 文本分类

9.2 文本匹配

9.3 序列标注

9.4 序列生成

9.1 文本分类

■ 文本分类

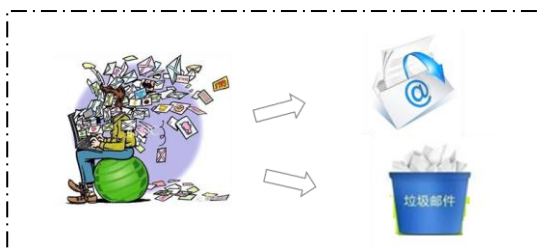
本节内容:

1. 分类任务概述
2. 序列结构文本分类
3. 图结构文本分类

1. 文本分类概述

■ 文本分类任务

利用计算机对大量的文档按照分类标准实现自动归类



分类任务	输入	任务建模	输出
	文本序列	分类模型	类别标签 类别标签根据任务定

$F: X \rightarrow Y$

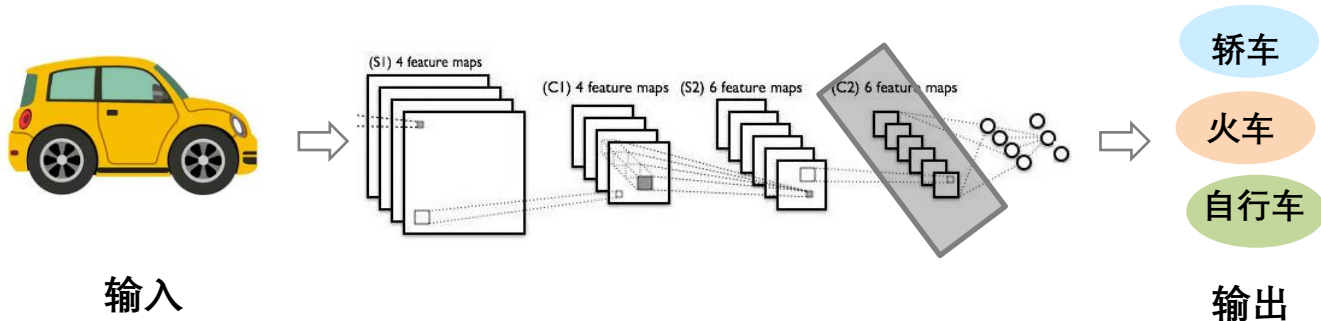
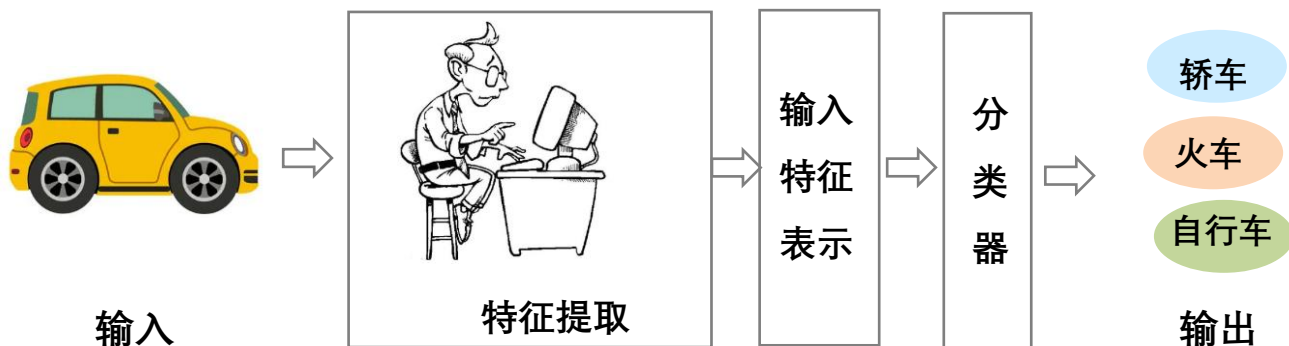
输入: X 句子/篇章

输出: X 所属类别 $Y \in \{\text{类别集合}\}$

1. 文本分类任务概述

■ 分类方法

概率统计时代：特征工程+算法（Naive Bayes/SVM/LR/KNN……）



深度学习时代：自动获取特征（表示学习）端到端分类

1. 文本分类概述

■ 神经网络分类方法

- ★ 基于词袋的文本分类
- ★ 基于卷积神经网络文本分类
- ★ 基于循环神经网络文本分类
- ★ 基于attention机制文本分类
- ★ 基于图卷积神经网络文本分类

也可以根据问题需要将上述方法结合形成混合模型

篇章级一般采用层次化的方法，先得到句子编码，然后以句子编码为输入，进一步得到篇章的表示

9.1 文本分类

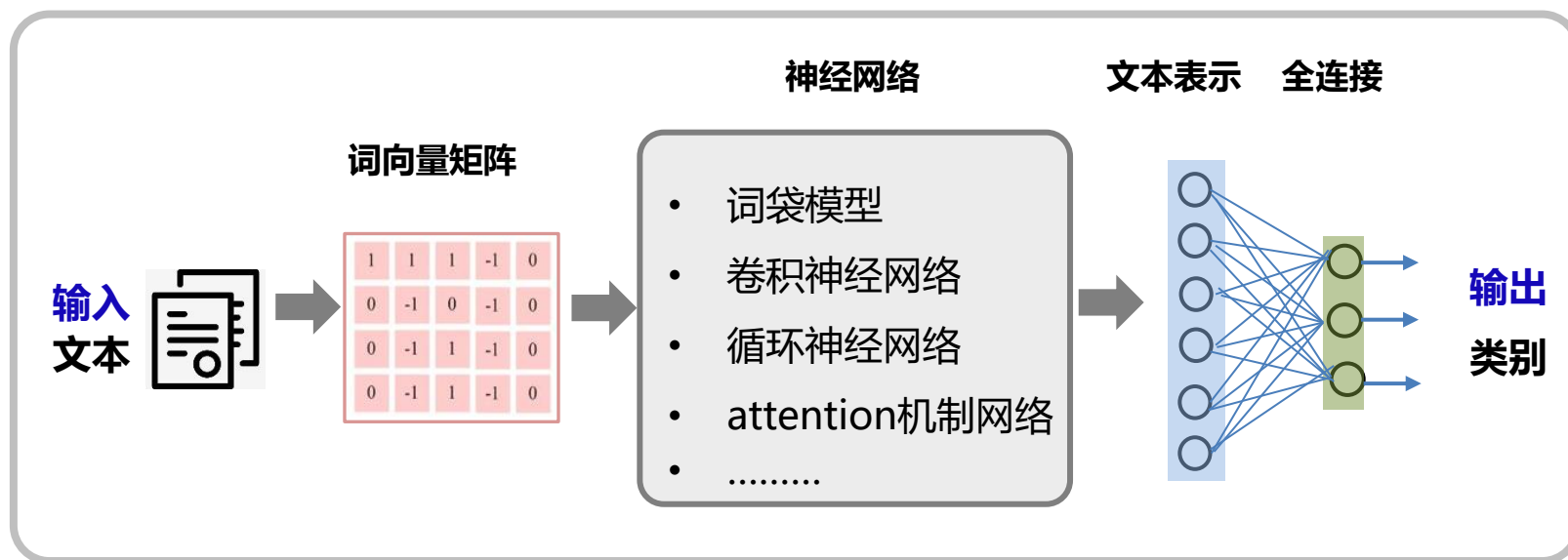
■ 文本分类

本节内容:

1. 分类任务概述
2. 序列结构文本分类
3. 图结构文本分类

2. 序列结构文本分类

■ 序列结构文本分类框架（文本整体分类）



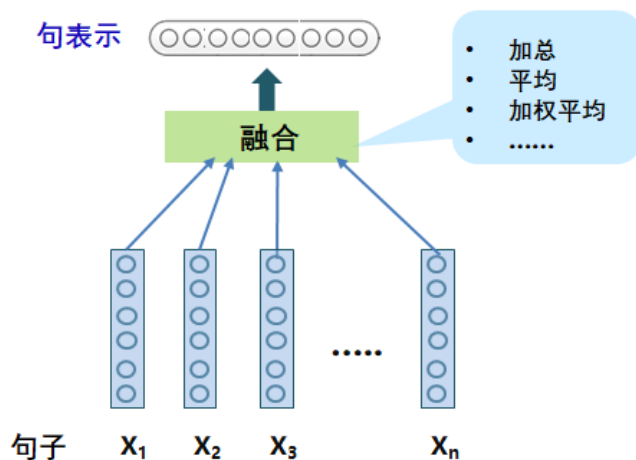
分类任务	输入	任务建模	输出
	文本序列	分类模型	类别标签 类别标签根据任务定

关键问题：如何生成高质量的文本表示

2. 序列结构文本分类

★ 神经词袋模型 (Bag of words)

简单对文本序列中每个词嵌入进行平均/加总，作为整个序列的表示。
这种方法的缺点是丢失了词序信息。对于长文本，神经词袋模型比较有效。但是对于短文本，神经词袋模型很难捕获语义组合信息



2. 序列结构文本分类

★ FastText

Facebook提出了一种简单而有效的文本分类和表示学习方法可以在不到10分钟的时间内使用标准的多核CPU对超过10亿个单词进行快速文本训练，并在不到一分钟的时间内对312K类中的50万个句子进行分类。

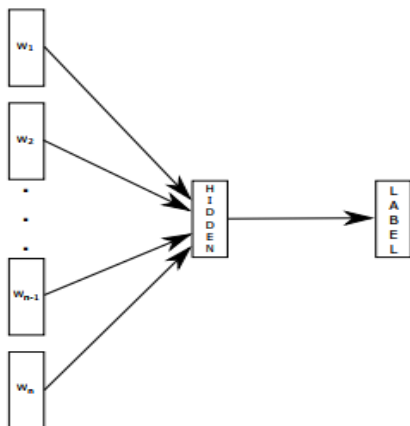


Figure 1: Model architecture for fast sentence classification.

输出层： 类别较少用softmax，
类别较多用hierarchical softmax

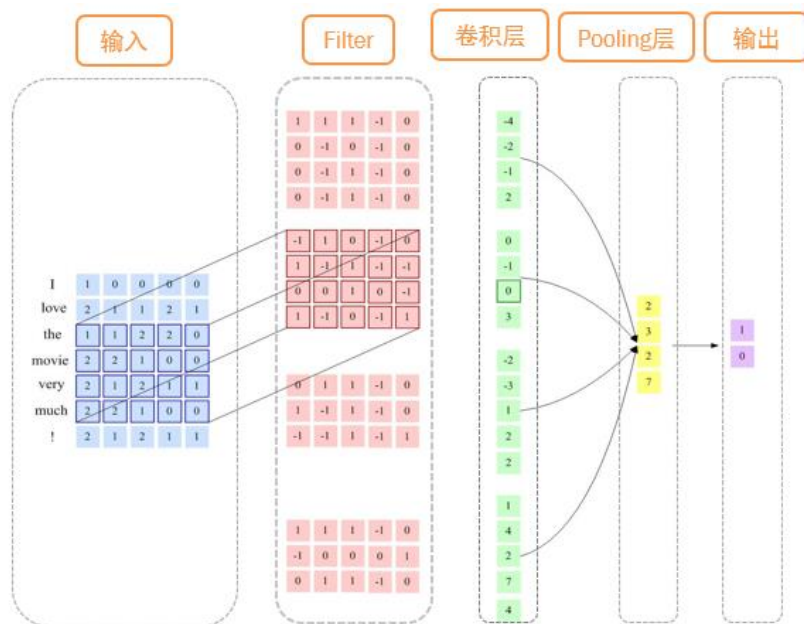
隐藏层： 将每个词向量相加取平均值

输入： Document中的每个词的词向量

2. 序列结构文本分类

★ 卷积神经网络模型 (Convolutional Neural Network)

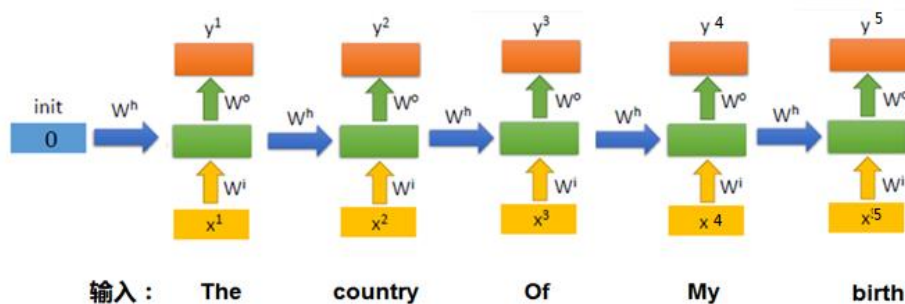
通过多个卷积层和子采样层，抽取序列的 n-gram特征信息，最终将得到特征信息合并成一个固定长度的向量作为整个序列表示。



2. 序列结构文本分类

★ 循环神经网络 (Recurrent Neural Network)

将文本序列看作时间序列，不断更新，最后得到整个序列的表示。这种表示中包含的是序列的顺序信息。

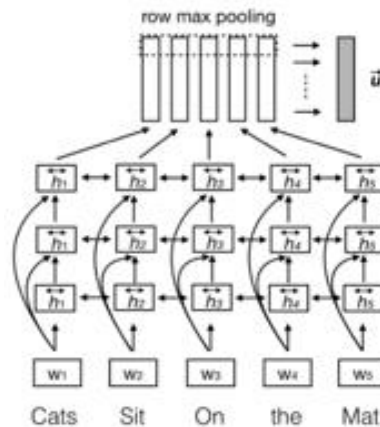
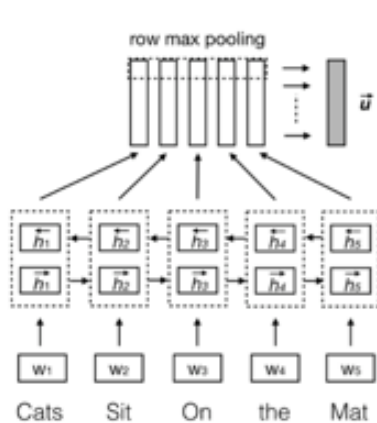
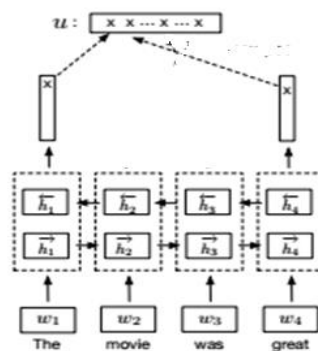
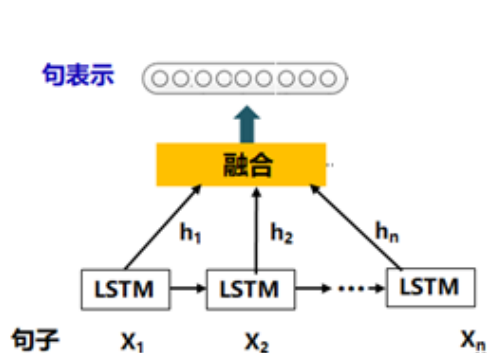


问题：RNN属于“biased model”，一个句子中越往后的词重要性越高，这有可能影响最后的分类结果，因为对句子分类影响最大的词可能处在句子任何位置。

2. 序列结构文本分类

★ 基于RNN的句表示

可根据需要形成各种基于RNN的句表示



2. 序列结构文本分类

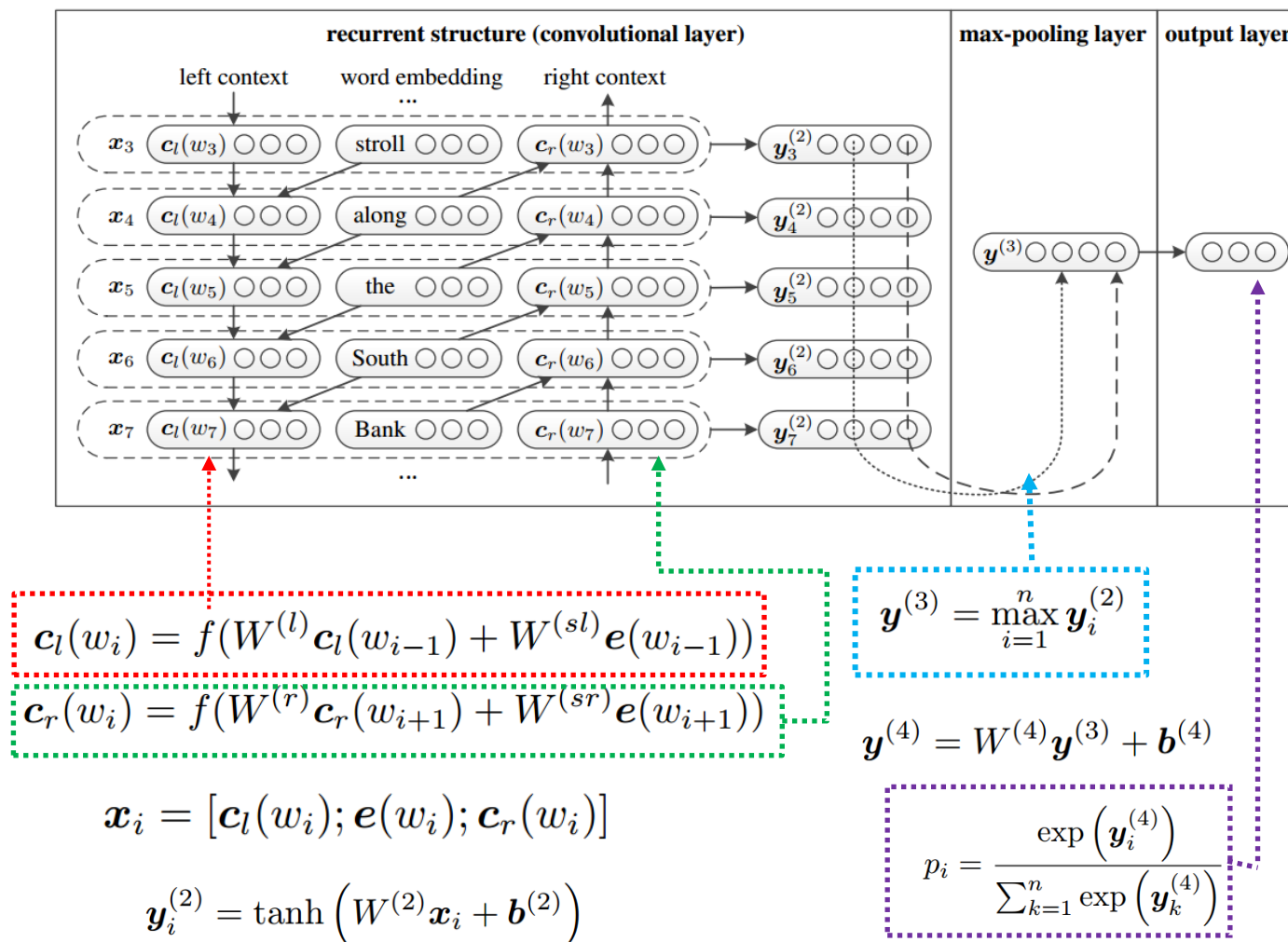
★ 循环+卷积神经网络模型

- RNN擅长处理序列结构，能够考虑到句子的上下文信息。
- CNN属于无偏模型，能够通过最大池化获得最重要的特征。

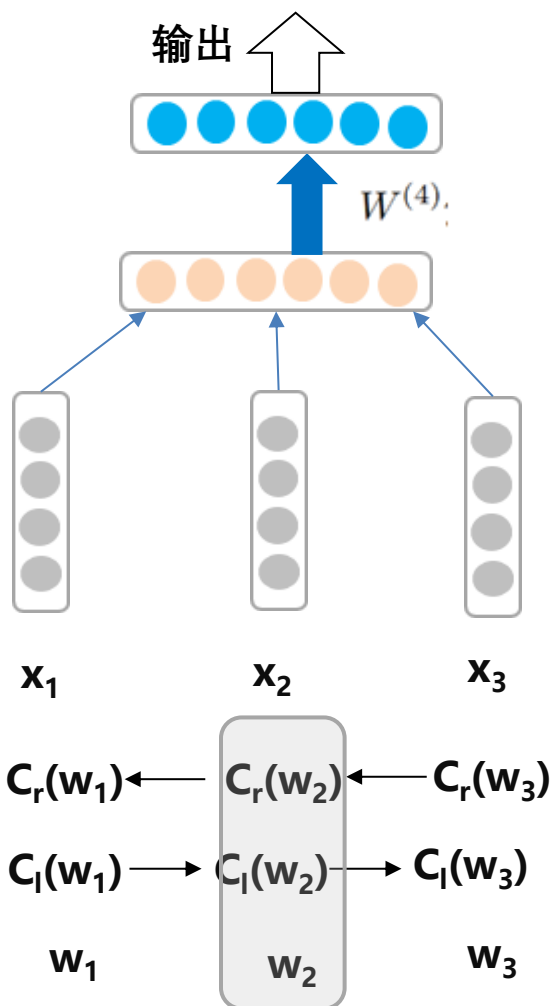
结合二者的优势生成上下文窗口信息的卷积网络

2. 序列结构文本分类

★ Recurrent Convolutional Neural Networks



2. 序列结构文本分类



$$p_i = \frac{\exp(y_i^{(4)})}{\sum_{k=1}^n \exp(y_k^{(4)})}$$

$$y^{(4)} = W^{(4)} y^{(3)} + b^{(4)}$$

$$y^{(3)} = \max_{i=1}^n y_i^{(2)}$$

$$\dots y_i^{(2)} = \tanh(W^{(2)} x_i + b^{(2)})$$

$$\dots x_i = [c_l(w_i); e(w_i); c_r(w_i)]$$

$$\dots c_r(w_i) = f(W^{(r)} c_r(w_{i+1}) + W^{(sr)} e(w_{i+1}))$$

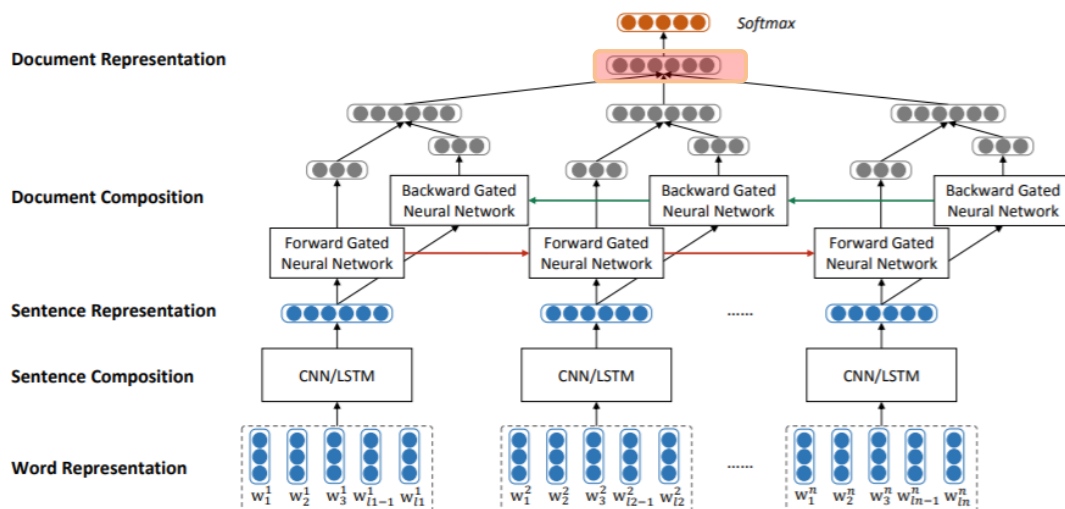
$$\dots c_l(w_i) = f(W^{(l)} c_l(w_{i-1}) + W^{(sl)} e(w_{i-1}))$$

.....

2. 序列结构文本分类

★ LSTM/CNN-GRU (篇章级-混合模型)

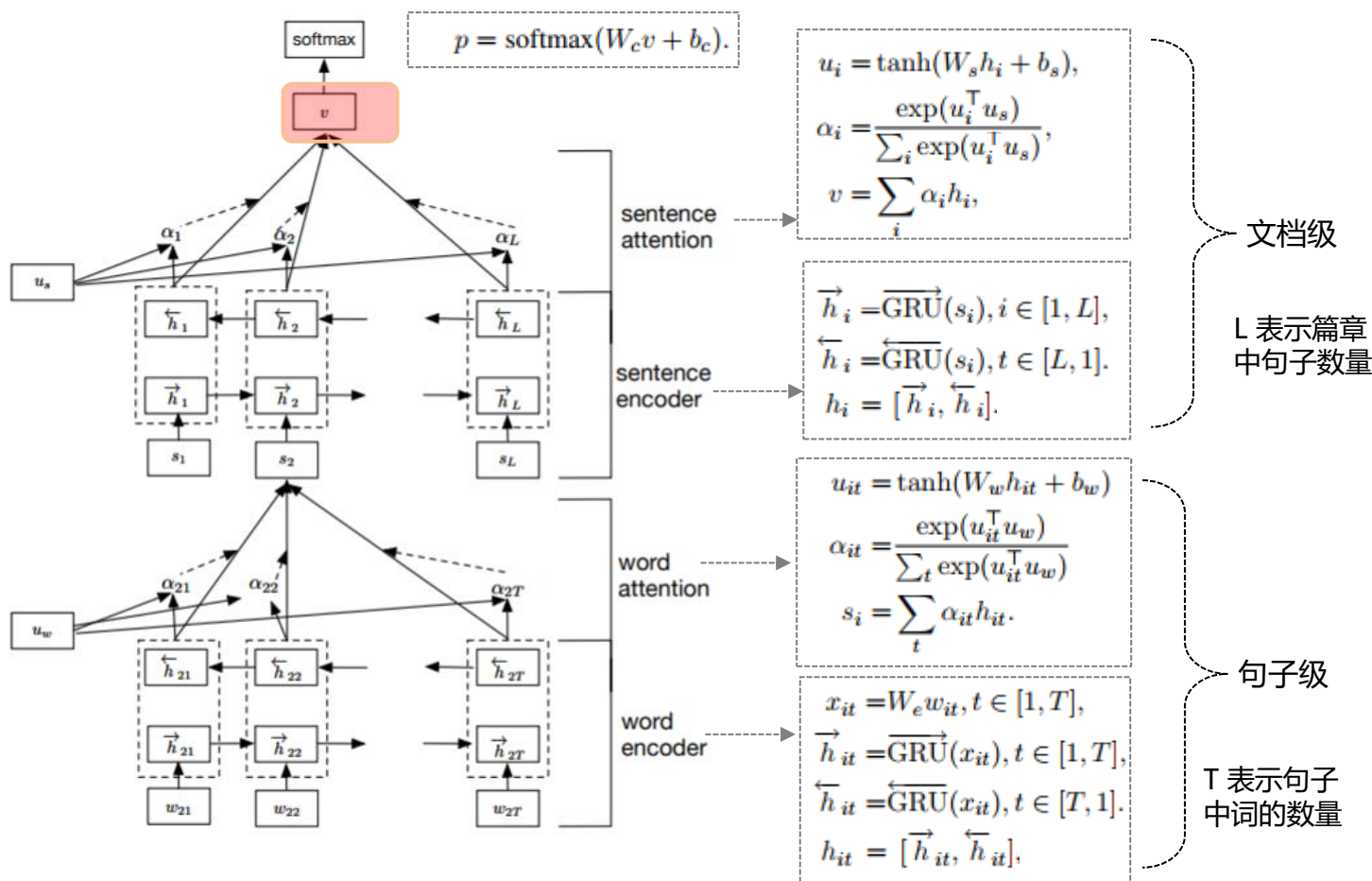
- 篇章中所有句子的词向量矩阵作为输入
- 用CNN/LSTM形成句子级向量表示
- 由句向量用双向RNN 形成每句的带有上下句子信息的句子表示
- 由句向量形成篇章级向量表示
- 用篇章级向量做分类



2. 序列结构文本分类

★ HAN (篇章级- Attention 模型)

从句子级和文档级两个层次引入Attention机制，可识别分类决策的重要单词和句子



2. 序列结构文本分类

■ 序列结构文本分类框架（有侧重点情况）

实体间关系抽取

例：比尔·盖茨作为微软公司的创始人

如，已知 实体“比尔·盖茨”和“微软公司”确定二实体之间的关系

输入：比尔·盖茨作为微软公司的创始人 能否只将二个实体做分类模型输入？

输出：二者关系 CEO 关系 \in {关系类别集合} - 限定关系抽取

问题：类似分类问题（非文本整体分类）如何建模？

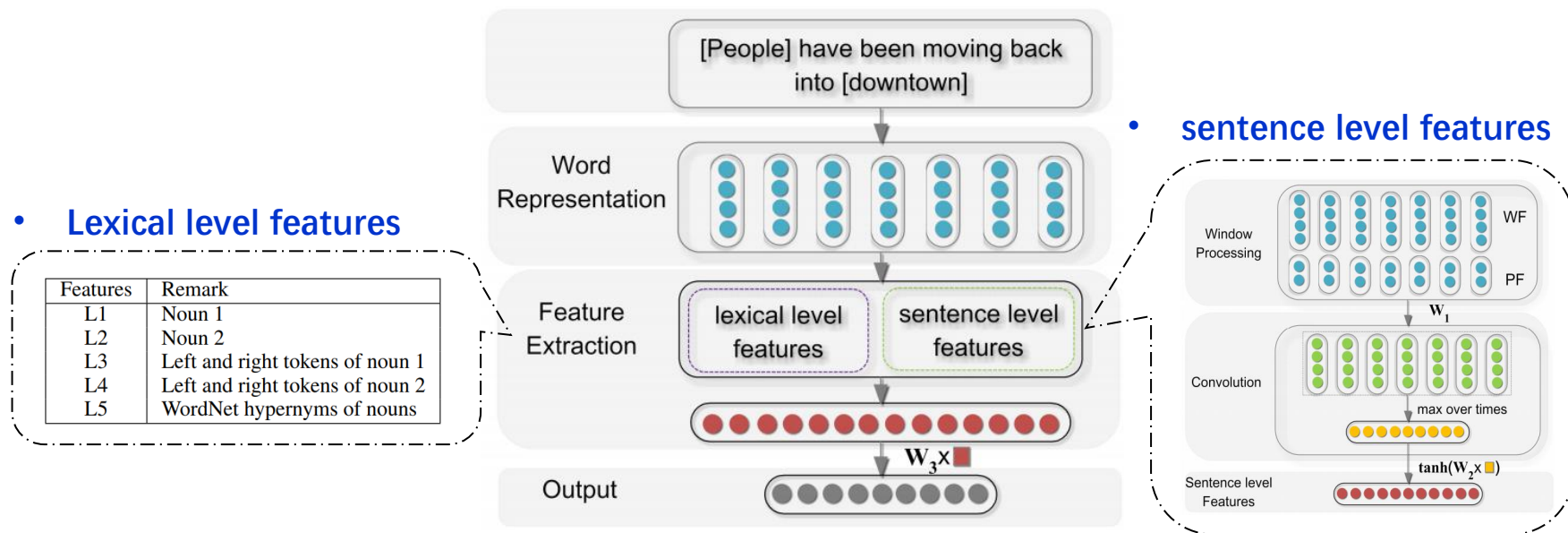
答：需要加入侧重点信息，增强模型性能

2. 序列结构文本分类

★CNN关系抽取 (2014)

论文主要是在输入时加入了一些Lexical Level以及position的特征：

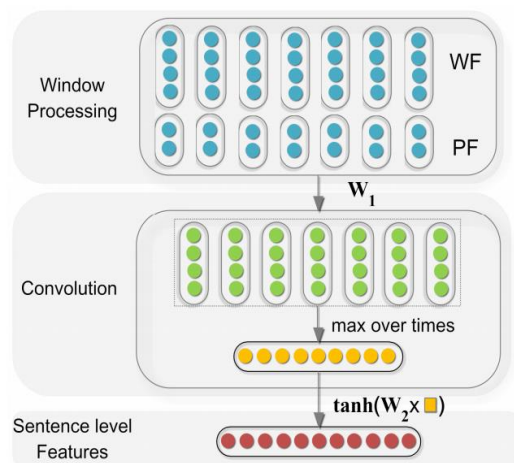
◆ 模型结构



输入：句子

输出：句子中实体之间的关系

2. 序列结构文本分类



- sentence level features

设 WF窗口 为 3

WF: $[X_s, X_0, X_1], [X_0, X_1, X_2], \dots, [X_5, X_6, X_e]\}$

PF是句子中每个词到两个实体的距离[d1, d2].

$S: [\text{People}]_0 \text{ have}_1 \text{ been}_2 \text{ moving}_3 \text{ back}_4 \text{ into}_5 [\text{downtown}]_6$

PF Been: $[2, -4]$ $[\text{WF}, \text{PF}]^T$

将上述特征拼接输入CNN后得到句子级特征。

将 Lexical level features 和 sentence level features 融入句表示
用融合后的句表示做分类

9.1 文本分类

■ 文本分类

本节内容:

1. 分类任务概述
2. 序列结构文本分类
3. 图结构文本分类

3. 图结构文本分类

■ 图卷积神经网络文本分类

根据任务对原文本加入附加信息并构建原文本与附加信息的关系图（将附加的结构信息融入文本），然后利用图卷积的方法提取文本有效的特征表示

图卷积文本分类步骤：

1. Graph 构建

对原文本按照附加信息的不同构建不同的图结构。附加信息可以是词的近义信息，共现信息，先验知识信息等

2. 文本Graph节点特征表示

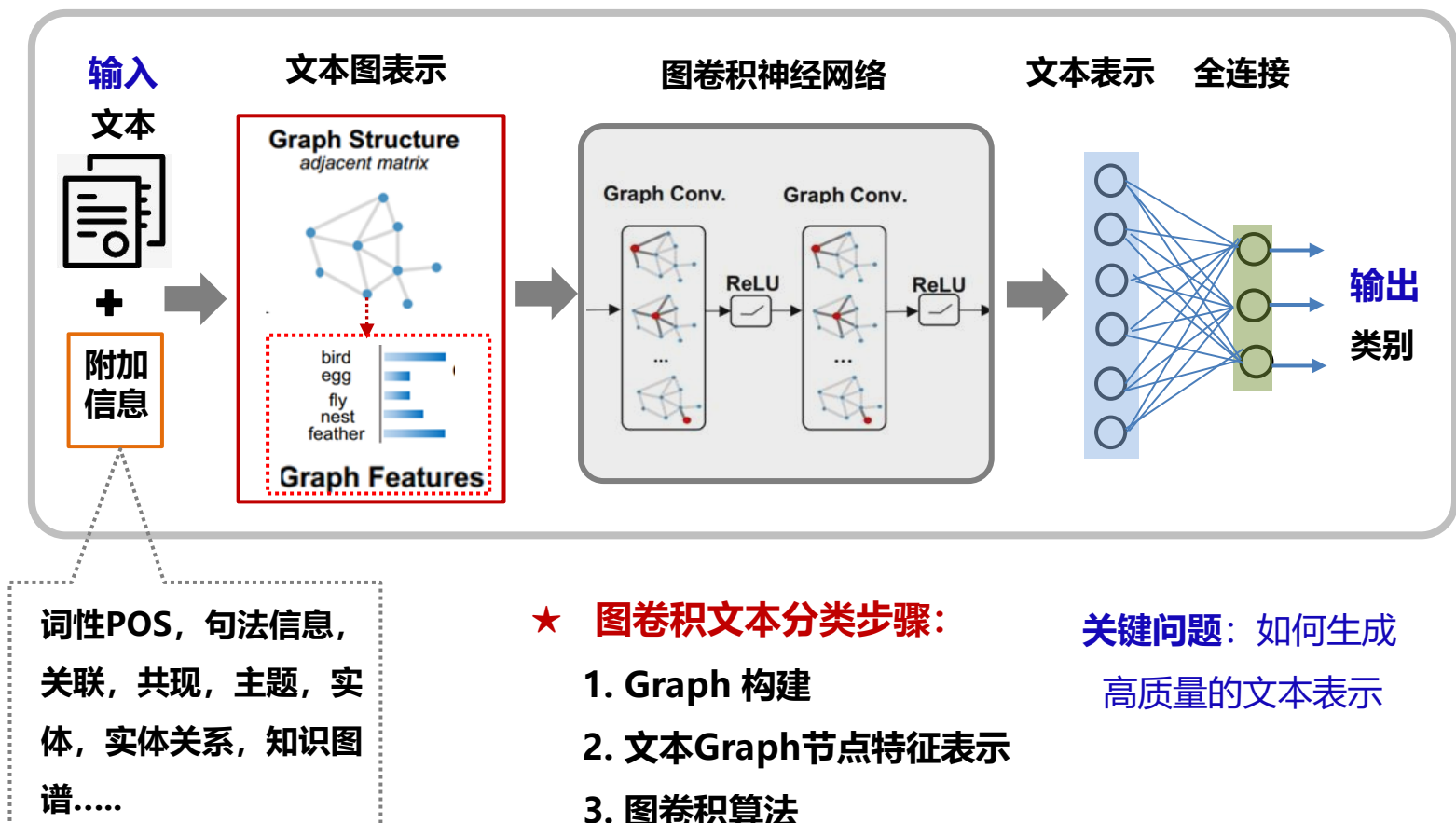
在图卷积中结点可以根据任务需要采用不同的结点表示方法。如，一般词向量，ELMO 词嵌入，Bi-LSTM 词向量嵌入，词袋词频等方法

3. 图卷积算法

构建好输入图和图上结点表示后，可以根据不同的任务构建不同的图卷积算法。如，一般图卷积，加入注意力机制的图卷积等

3. 图结构文本分类

■ 图文本分类框架



3. 图结构文本分类

例1. 对文本进行图卷积分类

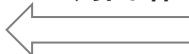
1. Graph 构建

文本

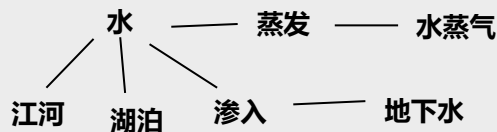
地面上的水不断蒸发变成了水蒸气，有的水在地面上汇成江河、湖泊，另外一些水渗入了地下，称为地下水。



查询图谱



知识图谱



地面上的水不断蒸发变成了水蒸气，有的水在地面上汇成江河、湖泊，另外一些水渗入了地下，称为地下水。



邻接表形式图结构

注：附加信息可以是词性POS，句法信息，关联，共现，主题，实体，实体关系，知识图谱 等

3. 图结构文本分类

2. 文本Graph节点特征表示

地面上的水不断蒸发变成了水蒸气，有的水在地面上汇成江河、湖泊，另外一些水渗入了地下，称为地下水。

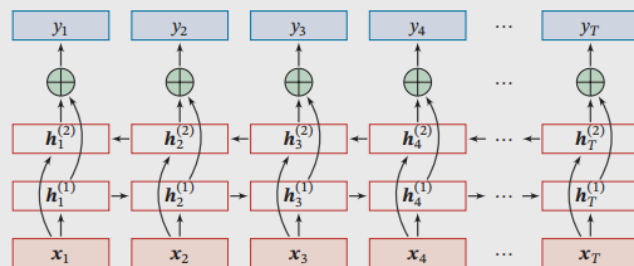
结点表示

表示方法

地面
水
不断
蒸发
变成
水蒸气
.....



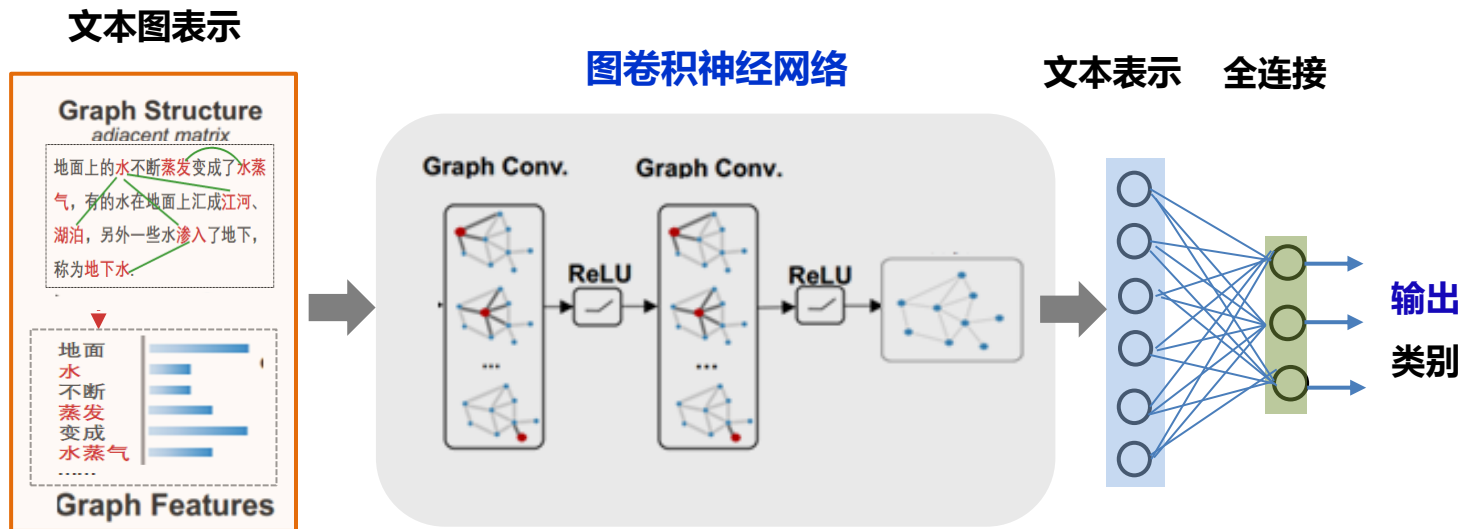
双向LSTM



注：结点表示可根据具体问题需要采用双向RNN，Elmo词嵌入等各种表示方法

3. 图结构文本分类

3. 图卷积算法



注：可以根据不同的任务构建不同的图卷积算法。如，图卷积网络文本分类算法，图卷积网络多层加权分类算法等

内 容 提 要

9.0 概述

9.1 文本分类

9.2 文本匹配

9.3 序列标注

9.4 序列生成

9.2 文本匹配

■ 文本匹配

本节内容：

1. 文本匹配任务概述
2. 孪生网络方法
3. 交互耦合方法

1. 文本匹配任务概述

■ 文本匹配

广义的将研究两段文本间关系的问题定义为“**文本匹配**”问题，**匹配**含义根据任务的不同有不同的定义。很多自然语言处理的任务都会涉及文本匹配问题

- 复述识别 (paraphrase identification)

两个句子“**感冒了是否要吃药**”和“**感冒了要吃什么药**”

问：两个句子是否表达同样的意思？

- 文本蕴含识别 (Textual Entailment)

两个句子“**我正在上海旅游**”和“**我正在八达岭长城**”

问：这两句话是什么关系？

匹配任务	输入	任务建模	输出
	二段文本序列	匹配模型	二者关系 一般类别标签

1. 文本匹配任务概述

■ 与文本匹配相关的NLP任务

Tasks	Text 1	Text 2	Objective
Paraphrase Identification	string 1	string 2	classification
Textual Entailment	text	hypothesis	classification
Question Answer	question	answer	classification/ranking
Conversation	dialog	response	classification/ranking
Information Retrieval	query	document	ranking

https://blog.csdn.net/Ding_xiaofei

这些任务的共性问题是研究两段文本之间关系的问题，一般可建模为“**分类**”和“**排位**”问题

1. 文本匹配任务概述

1.复述识别 (paraphrase identification)

又称释义识别，是判断两段文本是不是表达了同样的语义，这一类场景一般建模成分类问题。

如：两个句子“感冒了是否要吃药”和“感冒了要吃什么药” 问：两个句子是否表达同样的意思？

解决方法：该问题的句子匹配是计算二个句子相似度，可建模为二分类问题

1. 文本匹配任务概述

2. 文本蕴含识别 (Textual Entailment)

给定一个前提文本 (text)，根据这个前提去推断假说文本 (hypothesis) 与文本的关系，关系有：蕴含关系 (entailment)，矛盾关系 (contradiction)，蕴含关系 (entailment)。这一类场景一般建模成多分类问题。

如：两个句子“我正在上海旅游”和“我正在八达岭长城” 问：这两句话是什么关系？

解决方法：该问题的句子匹配是计算二个句子之间的关系，可建模为多分类问题

1. 文本匹配任务概述

3.问答 (QA)

根据Question在段落或文档中查找Answer，这类场景常常会被建模成分类问题；还有一类是根据Question从若干候选中找出正确答案，这类场景常常会被建模成排位（ranking）问题。

4.对话 (Conversation)

与QA 类似，但是比QA更复杂，由于引入了历史轮对话，需要考虑在历史轮的限制下回复是否合理。一般建模为分类或排位问题。

5.信息检索 (IR)

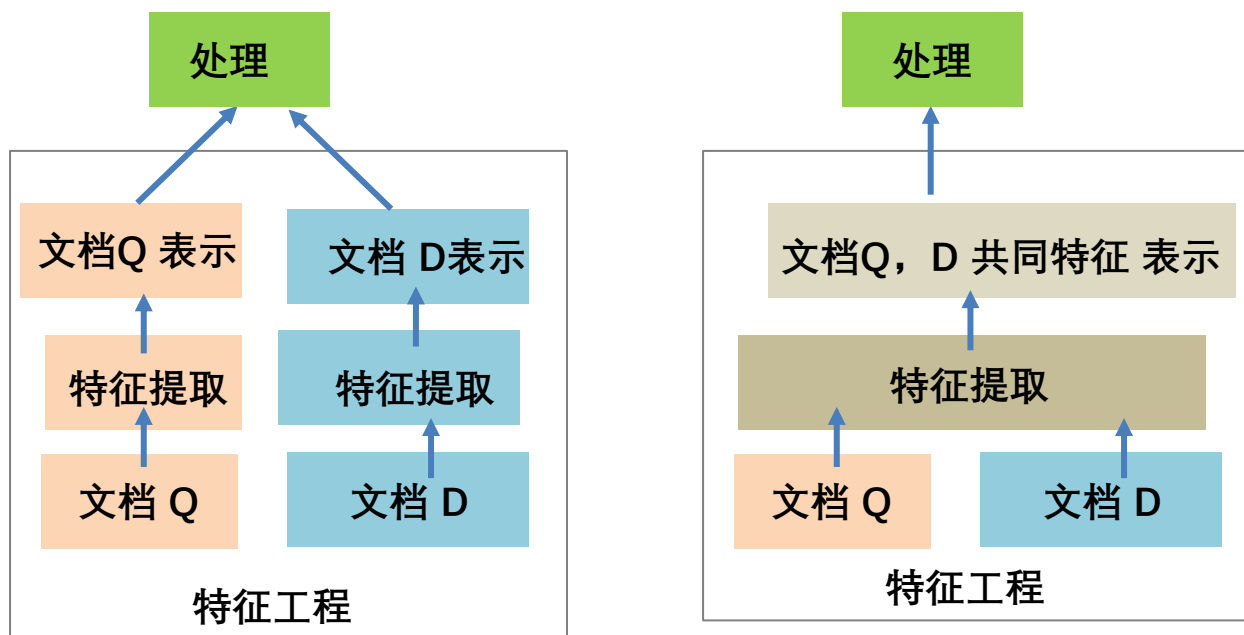
信息检索是一个更为复杂的任务，往往会有Query—Tittle，Query—Document的形式（Query可能是一个Document）检索需要计算相似度和排序一般建模为排位问题。

1. 文本匹配任务概述

■ 匹配方法

★ 统计方法：特征工程+算法（PRanking / margin/ SVM/LR……）

以上二种传统文本匹配方法主要集中在人工定义特征之上的关系学习，焦点在于如何人工提取的特征和设置合适的文本匹配学习算法来学习到最优的匹配模

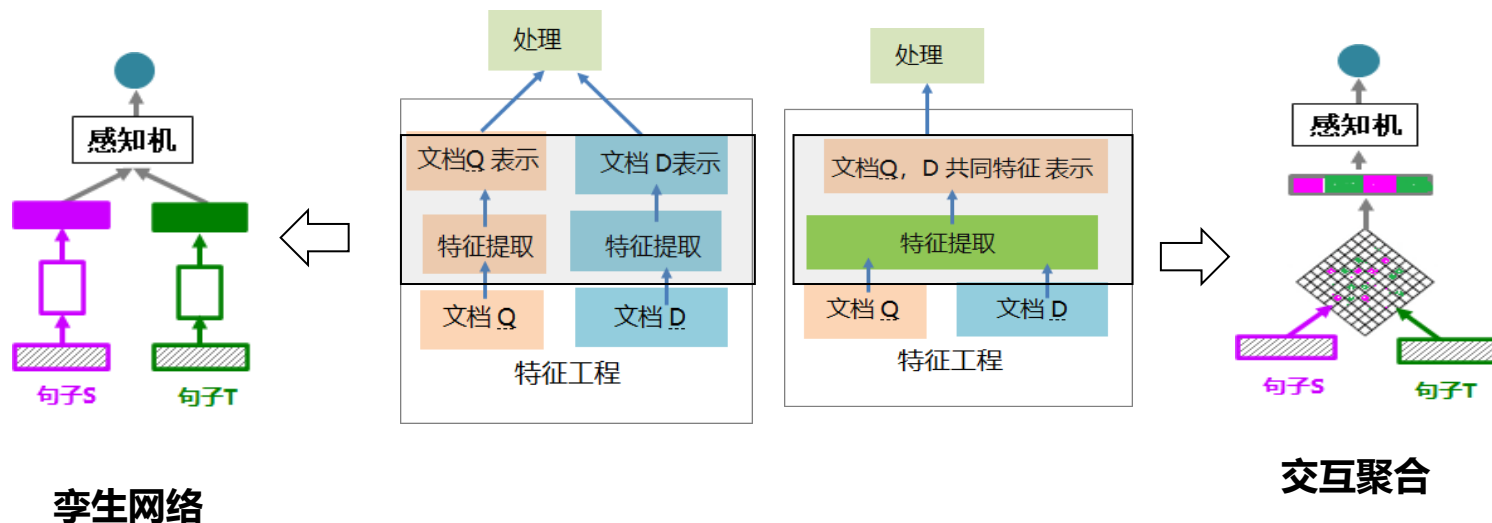


1. 文本匹配任务概述

★ 深度学习方法：

自动提取出词语之间的关系特征并结合短语匹配中的结构信息和文本匹配的层次化特性，更精细地描述文本匹配问题。

表示学习抽取有用特征



1. 文本匹配任务概述

匹配方法：

◆ 基于单语义文档表达的深度学习模型（基于表示-孪生网络）

主要思路： 首先将单个文本先表达成一个稠密向量（分布式表达）
然后直接计算两个向量间的相似度作为文本间的匹配度。

◆ 基于多语义文档表达的深度学习模型（基于交互-交互聚合）

主要思路： 需要建立多语义表达，更早地让两段文本进行交互，然后挖掘文本交互后的模式特征，综合得到文本间的匹配度。

复杂问题建模中常用交互耦合方法进行序列间的耦合表示，为后继挖掘交互表示模式特征提供信息。如选择式阅读理解等

9.2 文本匹配

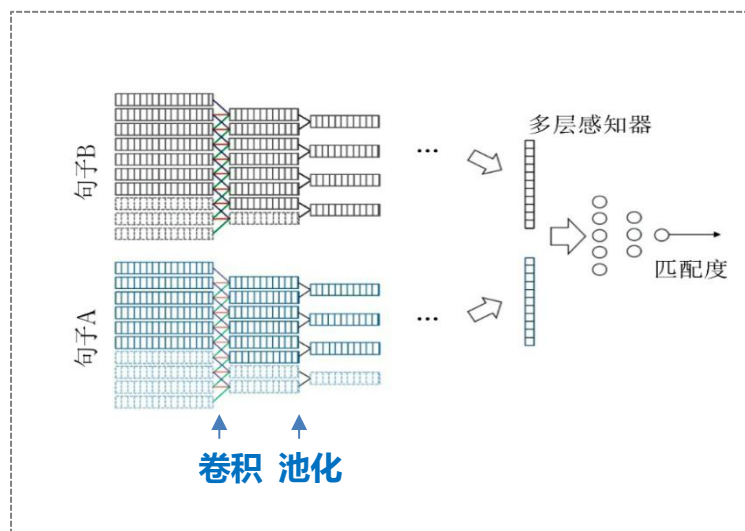
■ 文本匹配

本节内容：

1. 文本匹配任务概述
2. 孪生网络方法
3. 交互耦合方法

2. 孪生网络方法

★ ARC-I (基于CNN)



输入：句子A和B

运算关系：

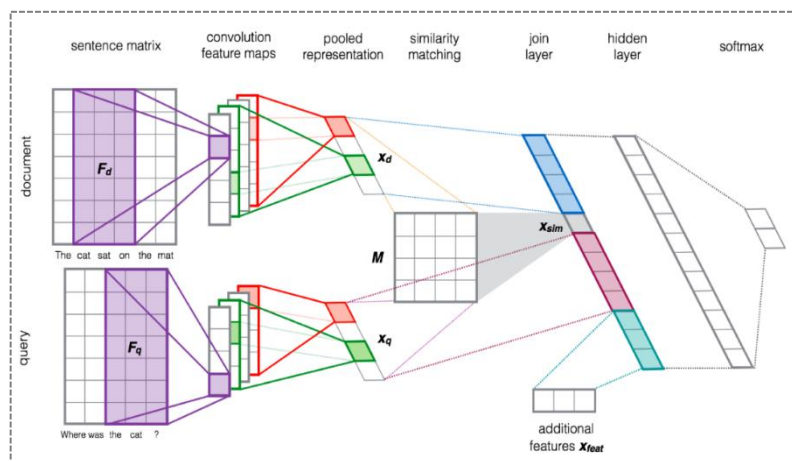
1. padding句子A/B到定长
2. 分别做多轮的卷积+池化运算
3. 拼接两个向量
4. 输入给多层感知机

输出：句子A和B的匹配度

特点： 在于将两个句子encode成句向量之后再用多层感知机进行分类，
没有体现出句子之间的交互操作

2. 孪生网络方法

★ CDNN (基于CNN)



特点：引入相似矩阵

输入：句子d和q

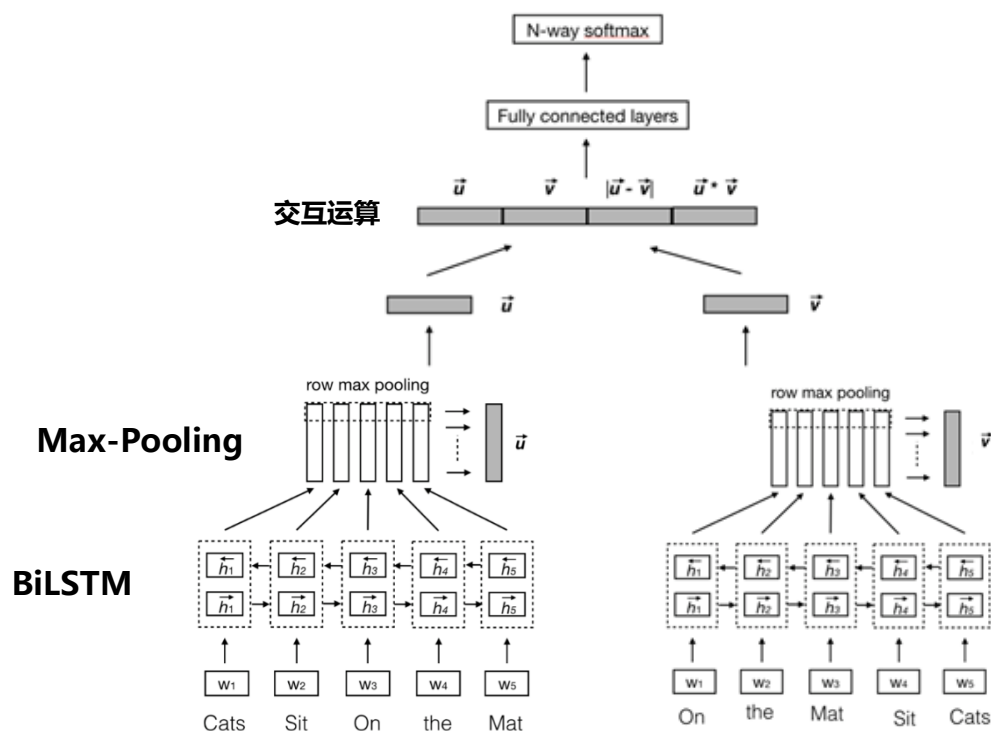
运算关系：

1. 分别做卷积+池化运算，得到句向量
2. 计算相似度 $X_{sim} = X_q^T M X_d$
3. 计算单词重叠数等其他特征
4. 拼接句向量、相似度和其他特征
5. 输入给多层感知机

输出：句子d和q的匹配度

2. 孪生网络方法

★ InferSent (基于RNN)



输入：句子A和B

运算关系：

1. 将A，B分别通过BiLSTM-Max表示成句向量
2. 将2个句向量交互运算输入给多层感知机

输出：句子A和B的匹配度

9.2 文本匹配

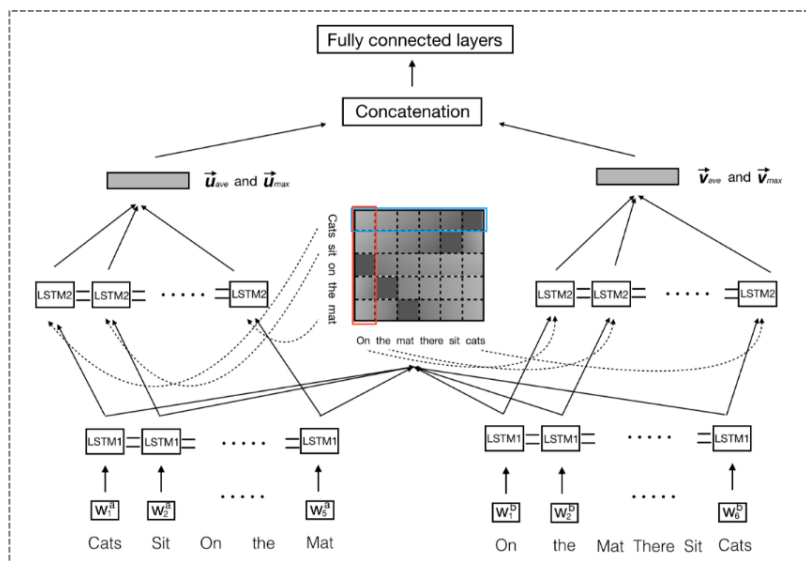
■ 文本匹配

本节内容:

1. 文本匹配任务概述
2. 孪生网络方法
3. 交互耦合方法

2. 交互聚合方法

★ ESIM (基于RNN 和 Attention)



运算关系：

1. 把句子X/Y表达成词向量序列，输入第一层BiLSTM进行编码
2. 将得到的词表示相互计算Attention得到交互矩阵
3. 将权重化的词表示与Attention值进行交互计算，再输入第二层BiLSTM
4. 将Max-Pooling和Aver-Pooling的结果串联
5. 输入给多层感知机

输入：句子X和Y

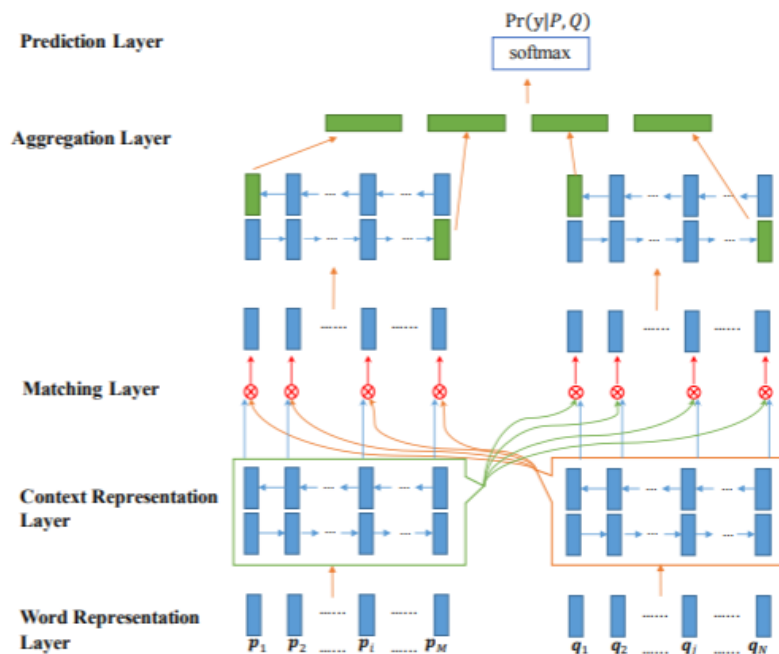
输出：句子X和Y的匹配度

优势：Attention机制更好地建模句子关系

2. 交互聚合方法

★ BiMPM(基于RNN)

该模型主要用于做文本匹配，即计算文本相似度。创新点在于采用了双向多角度匹配，采用matching-aggregation的结构，把两个句子之间的单元做相似度计算，最后经过全连接层与softmax层得到最终的结果



ESIM模型包含五部分：

- Word Representation Layer
- Context Representation Layer
- Matching Layer
- Aggregation Layer
- Prediction Layer

2. 交互聚合方法

例：选择机器阅读理解

段落

工商协进会报告，12月消费者信心上升到78.1，明显高于11月的72。另据《华尔街日报》报道，2013年是1995年以来美国股市表现最好的一年。这一年里，投资美国股市的明智做法是追着“傻钱”跑。所谓的“傻钱”策略，其实就是买入并持有美国股票这样的普通组合。这个策略要比对冲基金和其它专业投资者使用的更为复杂的投资方法效果好得多。

问题1：什么是傻钱策略？

答案：买入并持有美国股票这样的普通组合

问题2：12月的消费者信心指数是多少？

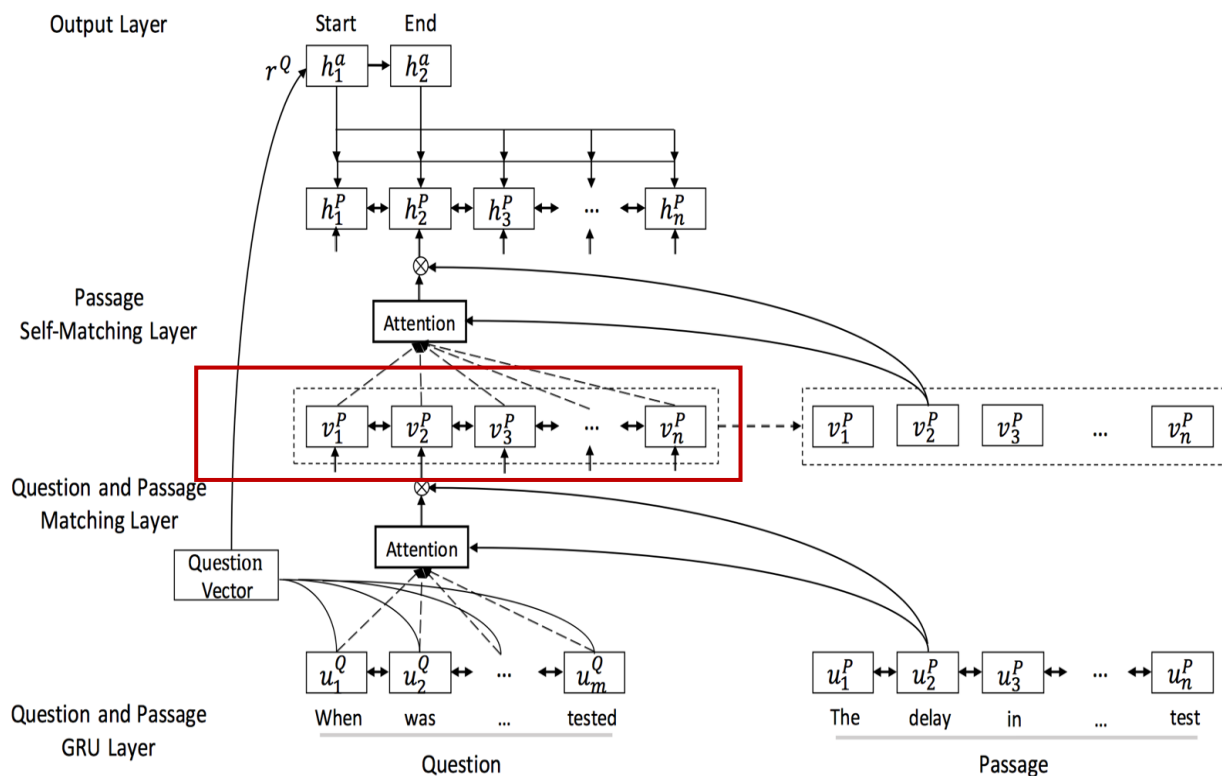
答案：78.1

问题3：消费者信心指数由什么机构发布？

答案：工商协进会

2. 交互聚合方法

R-NET 模型



交互耦合表示 (匹配)

输入: 文章+问题

参考文献:

- grave et, al. Bag of Tricks for Efficient Text Classification, 2017
- Kim et, al. Convolutional Neural Networks for Sentence Classification, 2014
- Lai et al. Recurrent Convolutional Neural Networks for Text Classification, 2015
- Yang et al. Hierarchical Attention Networks for Document Classification., 2016
- Hu et al. Heterogeneous Graph Attention Networks for Semi-supervised Short Text Classification 2019
- <https://www.jiqizhixin.com/articles/2019-10-18-14>
- <https://www.cnblogs.com/ZhangHT97/p/13391689.html>
- Alexis Conneau, Douwe Kiela, Holger Schwenk, Loic Barrault, and Antoine Bordes. 2017. Supervised learning of universal sentence representations from natural language inference data. EMNLP 2018.
- Qian Chen, Xiaodan Zhu, Zhenhua Ling, Si Wei, Hui Jiang, and Diana Inkpen. Enhanced LSTM for natural language inference. ACL 2017.
- Devlin, Jacob, Ming-Wei Chang, Kenton Lee and Kristina Toutanova. BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. NAACL-HLT, 2018.
- https://blog.csdn.net/weixin_38526306/article/details/88425997

在此表示感谢!



中国科学院大学
University of Chinese Academy of Sciences

谢谢！

Thank you

