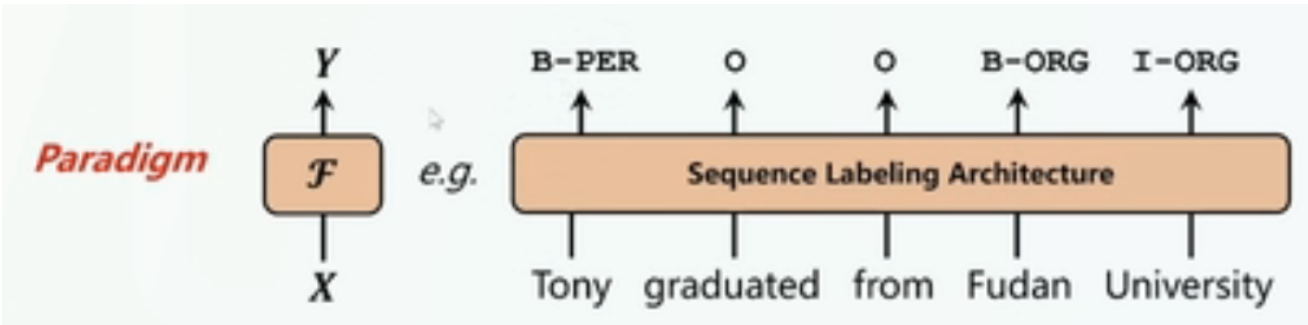


Paradigm Shift in Natural Language Processing

原文

范式

解决一类NLP任务的通用的框架



NLP七大范式

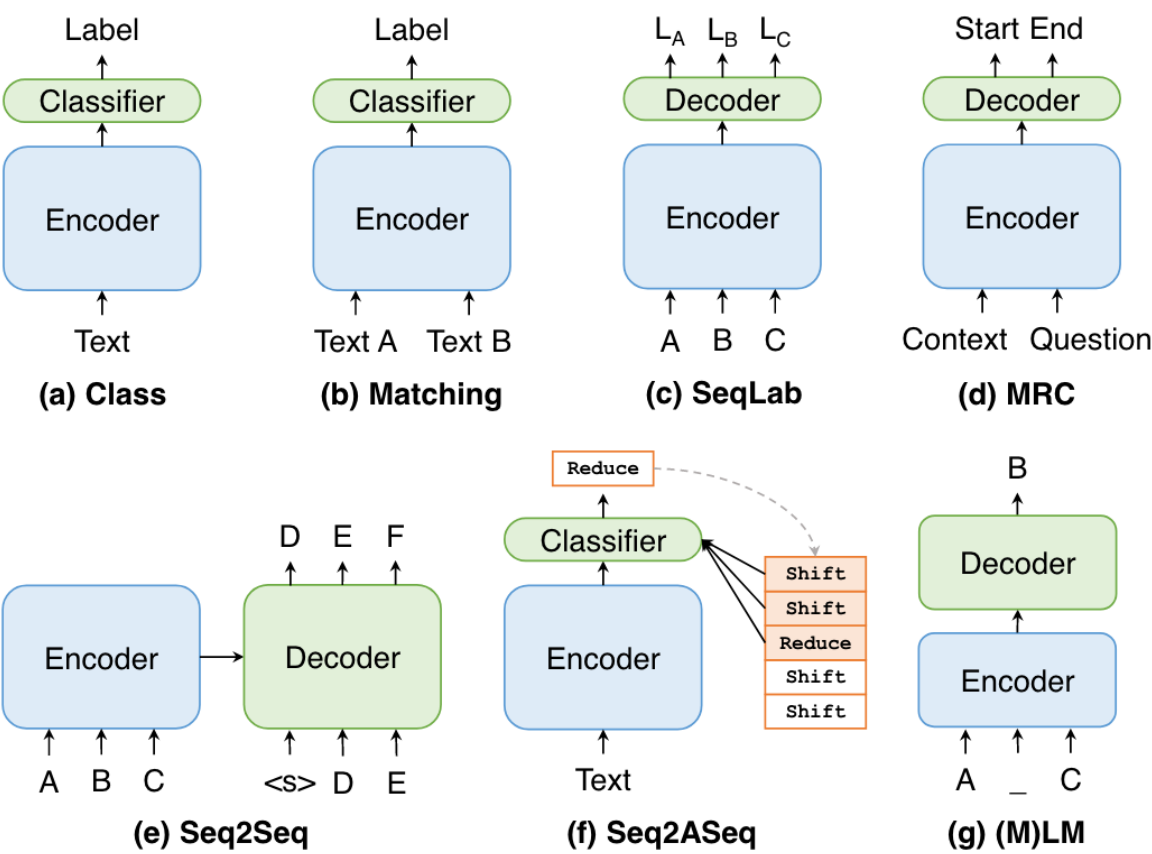
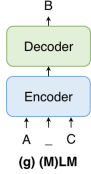


Figure 1: Illustration of the seven mainstream paradigms in NLP.

	图示		ENC	CLS/DEC	适用任务	备注
Class	 <p>(a) Class</p>	$\mathcal{Y} = \text{CLS}(\text{ENC}(\mathcal{X}))$	CNN/RNN/Transformer	(avg/max/attention) pooling + MLP	情感分析、主题分类、垃圾邮件检测....	
Matching	 <p>(b) Matching</p>	$\mathcal{Y} = \text{CLS}(\text{ENC}(\mathcal{X}_a, \mathcal{X}_b))$ $\text{Enc} = \text{enc2}(\text{enc1}(x_1), \text{enc1}(x_2))$	对两个文本分开或联合编码	捕获文本对的关系，并预测	NLI, 句子相似度	
SeqLab(Sequence labeling)	 <p>(c) SeqLab</p>	$y_1, \dots, y_n = \text{DEC}(\text{ENC}(x_1, \dots, x_n))$	RNN/Transformer	CRF	NER, POS-Tagging	
MRC (machine reading comprehension)	 <p>(d) MRC</p>	$y_k, \dots, y_{k+l} = \text{DEC}(\text{ENC}(\mathcal{X}_c, \mathcal{X}_q))$	CNN/RNN/Transformer	start/end position prediction	MRC	根据 question，从 context 中抽取连续的 spans.
Seq2Seq	 <p>(e) Seq2Seq</p>	$y_1, \dots, y_m = \text{DEC}(\text{ENC}(x_1, \dots, x_n))$	CNN/RNN/Transformer	CNN/RNN/Transformer	机器翻译，对话	
Seq2Seq (sequence-to-action-sequence)	 <p>(f) Seq2Seq</p>	$\mathcal{A} = \text{CLS}(\text{ENC}(x), \mathcal{C})$	CNN/RNN/Transformer	基于输入文本和当前栈的状态预测 action	Dependency parsing	CLS除了 encoder 的输出，还要维护一个栈的状态

(M)LM		$\text{LM} : x_k = \text{DEC}(x_1, \dots, x_{k-1})$ $\text{MLM} : \bar{x} = \text{DEC}(\text{ENC}(\tilde{x}))$	CNN/RN N/Transf ormer	分类器或 自回归 decoder	LM/MLM	
-------	--	---	-----------------------------	------------------------	--------	--

组合范式

有的复杂任务可以通过多个范式的组合来解决

e.g. HotPotQA: matching + MRC

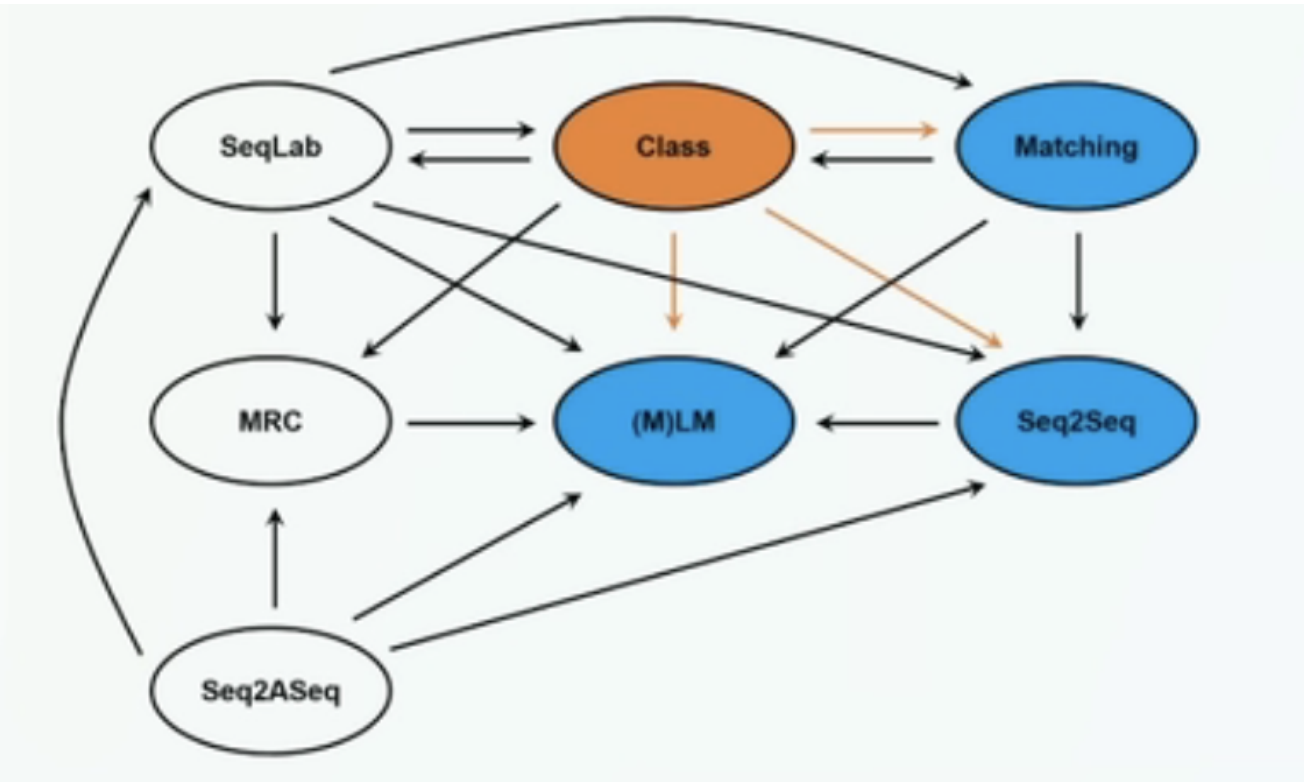
matching找到相关的document

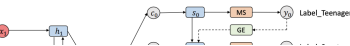
MRC从document中抽取出answer

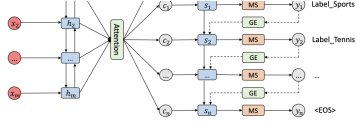
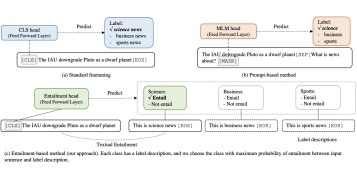
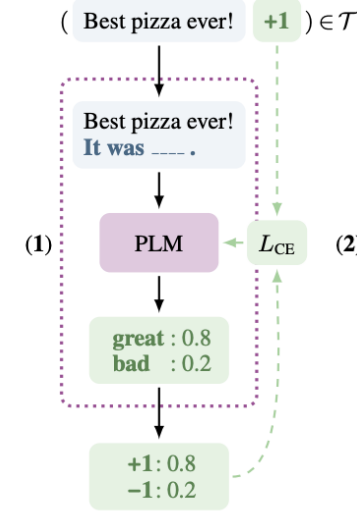
范式迁移

e.g.有的任务之前用Class解决，但也可以转化为(M)LM来解决

文本分类

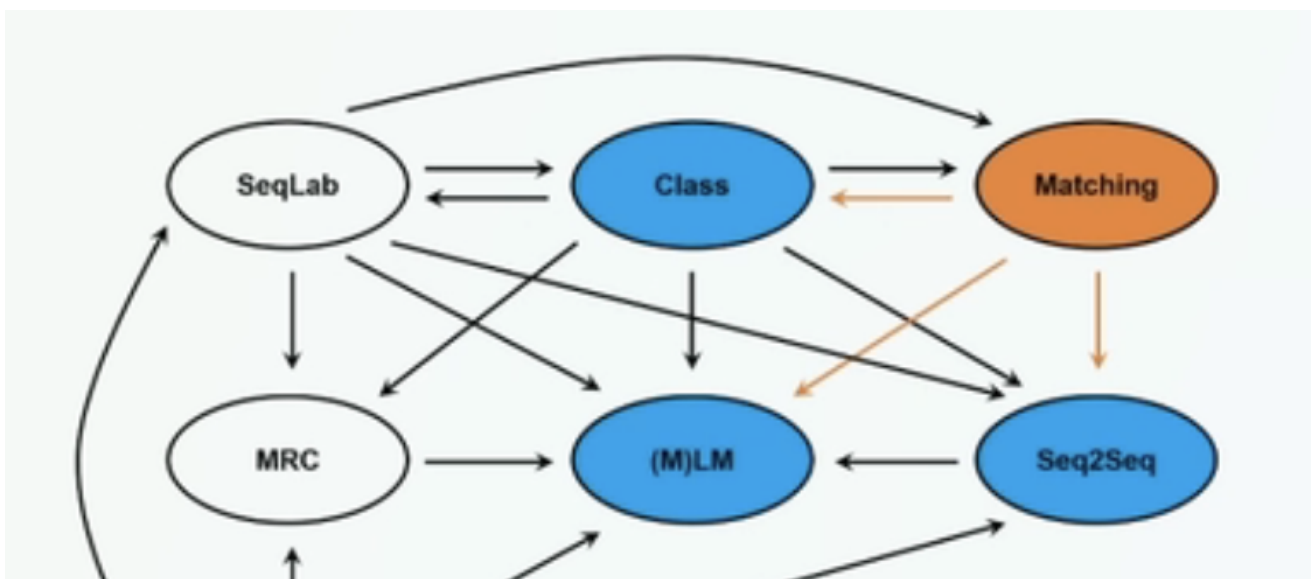


范式	图示	备注	论文
Seq2Se		多标签分类任务中，Class	SGM: sequence generation model

q		<p>没有考虑到标签之间的交互，Seq2Seq方法可以考虑之前预测的标签。</p>	<p>for multi-label classification. 2018 coling best paper</p>
Matchin g		<p>用原始句子和标签描述做一个NLI分类 (entail/not entail)</p>	<p>Entailment as few-shot learner.</p>
(M)LM			<p>Exploiting cloze-questions for few-shot text classification and natural language inference.</p>

NLI

自然语言推断任务，即给出一对句子, 判断两个句子是entailment/contradiction/neutral.



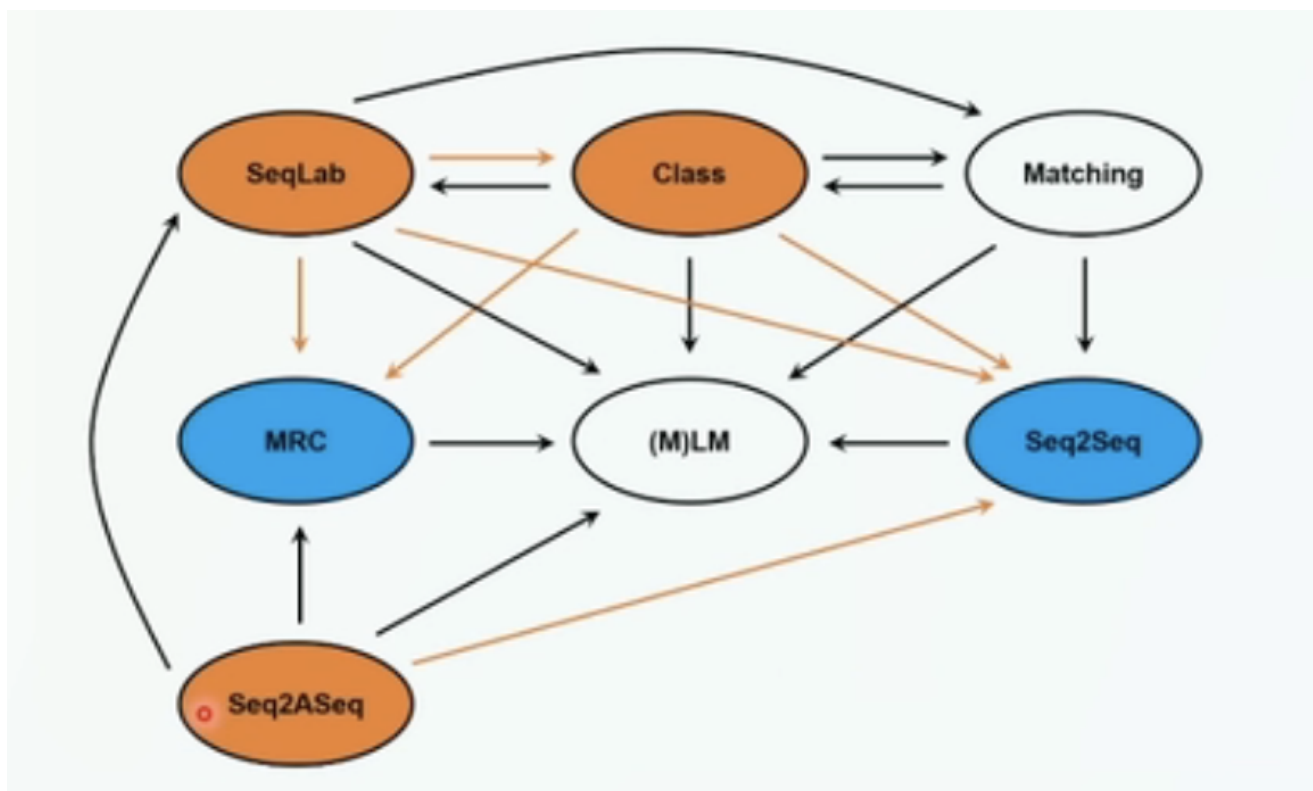
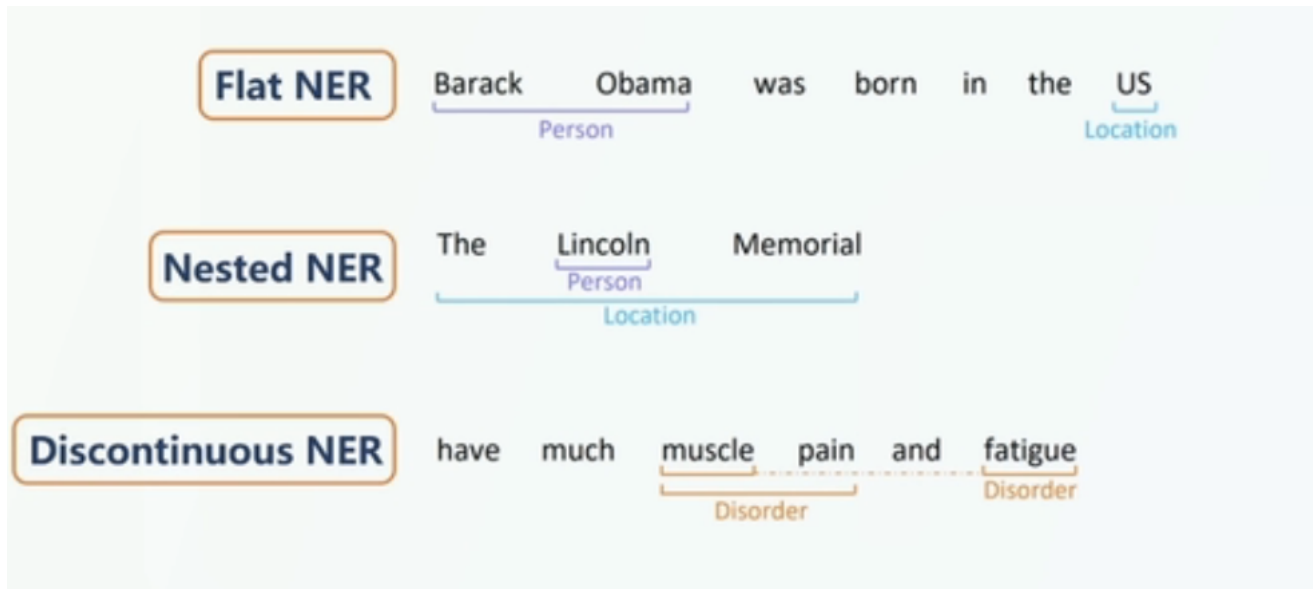


范式	图示	备注	论文
Class			
Seq2Seq		<p>Hypothesis: Product and geography are what make cream skimming work. Entailment, neutral, or contradiction?</p> <p>Premise: Conceptually cream skimming has two basic dimensions – product and geography. Entailment</p> <p>把所有任务都形式化为QA的形式，context, question, answer</p>	The Natural Language Decathlon: Multitask Learning as Question Answering
(M)LM	<p>MNLI The MNLI dataset (Williams et al., 2018) consists of text pairs $x = (a, b)$. The task is to find out whether a implies b (0), a and b contradict each other (1) or neither (2). We define</p> $P_1(x) = \text{"a"?} \parallel \text{---}, \text{"b"} \quad P_2(x) = \text{a?} \parallel \text{---}, b$		Exploiting cloze-questions for few-shot text classification and natural language inference.

and consider two different verbalizers v_1 and v_2 :

$v_1(0) = \text{Wrong}$ $v_1(1) = \text{Right}$ $v_1(2) = \text{Maybe}$
 $v_2(0) = \text{No}$ $v_2(1) = \text{Yes}$ $v_2(2) = \text{Maybe}$

NER

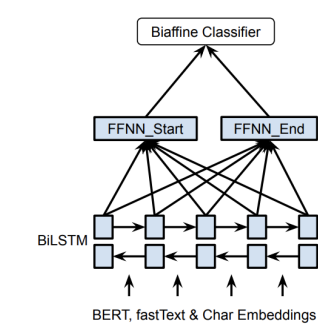


传统方法

SeqLab(flat NER)

Class(nested NER) 抽出所有的span，然后预测每个span属于什么实体

Seq2ASeq(discontinuous NER)

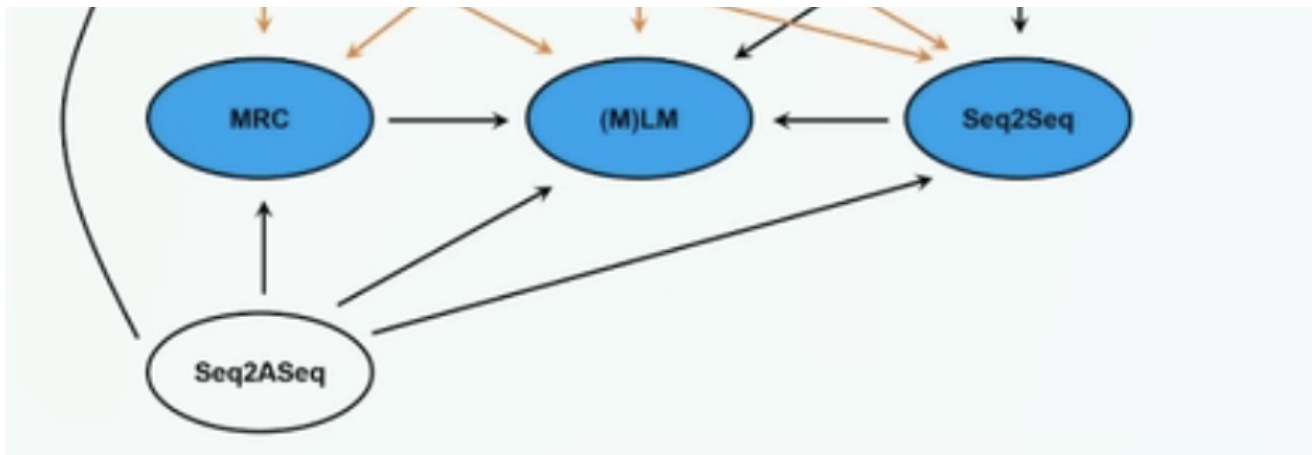
范式	图示	备注	论文												
Class(flat & nested)	<div><p>Figure 1: The network architectures of our system.</p><div><p>Matrix (l×l×c) Labeling:</p><table><tr><td>The</td><td>0</td><td>0</td><td>2</td></tr><tr><td>Lincoln</td><td>-1</td><td>1</td><td>0</td></tr><tr><td>Memorial</td><td>-1</td><td>-1</td><td>0</td></tr></table><p>The Lincoln Memorial</p>$h_s(i) = \text{FFNN}_s(x_{s_i})$$h_e(i) = \text{FFNN}_e(x_{e_i})$$r_m(i) = h_s(i)^T U_m h_e(i) + W_m(h_s(i) \oplus h_e(i)) + b_m$</div></div>	The	0	0	2	Lincoln	-1	1	0	Memorial	-1	-1	0	Biaffine: https://zhuanlan.zhihu.com/p/369851456	Named Entity Recognition as Dependency Parsing
The	0	0	2												
Lincoln	-1	1	0												
Memorial	-1	-1	0												
MRC(flat & Nested)	<table><tr><th>Entity</th><th>Natural Language Question</th></tr><tr><td>Location</td><td>Find locations in the text, including non-geographical locations, mountain ranges and bodies of water.</td></tr><tr><td>Facility</td><td>Find facilities in the text, including buildings, airports, highways and bridges.</td></tr><tr><td>Organization</td><td>Find organizations in the text, including companies, agencies and institutions.</td></tr></table> <p>Table 1: Examples for transforming different entity categories to question queries.</p>	Entity	Natural Language Question	Location	Find locations in the text, including non-geographical locations, mountain ranges and bodies of water.	Facility	Find facilities in the text, including buildings, airports, highways and bridges.	Organization	Find organizations in the text, including companies, agencies and institutions.		A unified MRC framework for named entity recognition				
Entity	Natural Language Question														
Location	Find locations in the text, including non-geographical locations, mountain ranges and bodies of water.														
Facility	Find facilities in the text, including buildings, airports, highways and bridges.														
Organization	Find organizations in the text, including companies, agencies and institutions.														
Seq2Seq(All)	<div><p>S1: B-Per Barack I-Per Obama O was O born O in O the B-Loc US Person Location</p><p>(a) Sequence labelling for flat NER</p><p>S2: The Lincoln Memorial the Lincoln Memorial the Lincoln Memorial Person Location</p><p>(b) Span-based classification for nested NER</p><p>S3: have much muscle pain and fatigue Disorder Disorder</p><p>(c) Transition-based method for discontinuous NER</p><p>S1: Barack Obama <Person> US <Location> S2: The Lincoln Memorial <Location> Lincoln <Person> S2: muscle pain <Disorder> muscle fatigue <Disorder></p><p>(d) A unified generative solution for all NER tasks</p></div>	BART+pointer network	A unified generative framework for various NER subtasks.												

任务定义：

Figure 1: Illustration of seven ABSA subtasks.

```

graph LR
    SeqLab((SeqLab)) <--> Class((Class))
    Class <--> Matching((Matching))
    SeqLab --> Matching
  
```

传统方法

SeqLab(AE, OE, AOE)

Class(ALSC)

范式	图示	备注	论文
Matching(ALSC)	<p>X: LOC1 is often considered the coolest area of London. Aspect: <i>Safety</i></p> <p>↓</p> <p>QA-M What do you think of the <i>safety</i> of LOC1? [X] NLI-M LOC1- <i>safety</i>. [X] QA-B The polarity of the aspect <i>safety</i> of LOC1 is positive. [X] NLI-B LOC1- <i>safety</i> - positive. [X]</p>		Utilizing BERT for Aspect-Based Sentiment Analysis via Constructing Auxiliary Sentence
MRC(All)	<p>Original training example:</p> <ul style="list-style-type: none"> • input text: The <i>ambience</i> was <i>nice</i>, but <i>service</i> was not so great. • annotations: (<i>ambience</i>, <i>nice</i>, <i>positive</i>), (<i>service</i>, <i>no so great</i>, <i>negative</i>) <p>↙ ↘</p> <p>Converted training example 1:</p> <ul style="list-style-type: none"> • query-1: Find the <i>aspect terms</i> in the text. • answer-1: <i>ambience</i>, <i>service</i> • query-2: Find the <i>sentiment polarity and opinion terms</i> for <i>ambience</i> in the text. • answer-2: (<i>nice</i>, <i>positive</i>) <p>Converted training example 2:</p> <ul style="list-style-type: none"> • query-1: Find the <i>aspect terms</i> in the text. • answer-1: <i>ambience</i>, <i>service</i> • query-2: Find the <i>sentiment polarity and opinion terms</i> for <i>service</i> in the text. • answer-2: (<i>not so great</i>, <i>negative</i>) <p>Figure 3: Dataset conversion.</p>		A joint training dual-mrc framework for aspect based sentiment analysis.
Seq2Seq(All)	<p>Tokens: The [wine list] is [interesting] and has [good value], but the [service] is [dreadful]</p> <p>Position indices: 0 1 2 3 4 5 6 7 8 9 10 11 12 13 14</p> <p>Labels: Positive, Positive, Positive</p>	BART + pointer network	A Unified Generative Framework for Aspect-Based Sentiment

	<table><tr><th>Subtask</th><th>Target Sequence</th></tr><tr><td><i>AE</i></td><td>1, 2, 12, 12, </s></td></tr><tr><td><i>OE</i></td><td>4, 4, 7, 8, 14, 14, </s></td></tr><tr><td><i>ALSC</i></td><td>1, 2, POS, </s></td></tr><tr><td></td><td>12, 12, POS, </s></td></tr><tr><td><i>AOE</i></td><td>1, 2, 4, 4, 7, 8, </s></td></tr><tr><td></td><td>12, 12, 14, 14, </s></td></tr><tr><td><i>AESC</i></td><td>1, 2, POS, 12, 12, NEG, </s></td></tr><tr><td><i>Pair</i></td><td>1, 2, 4, 4, 1, 2, 7, 8, 12, 12, 14, 14, </s></td></tr><tr><td><i>Triplet</i></td><td>1, 2, 4, 4, POS, 1, 2, 7, 8, POS, 12, 12, 14, 14, POS, </s></td></tr></table>	Subtask	Target Sequence	<i>AE</i>	1, 2, 12, 12, </s>	<i>OE</i>	4, 4, 7, 8, 14, 14, </s>	<i>ALSC</i>	1, 2, POS, </s>		12, 12, POS, </s>	<i>AOE</i>	1, 2, 4, 4, 7, 8, </s>		12, 12, 14, 14, </s>	<i>AESC</i>	1, 2, POS, 12, 12, NEG, </s>	<i>Pair</i>	1, 2, 4, 4, 1, 2, 7, 8, 12, 12, 14, 14, </s>	<i>Triplet</i>	1, 2, 4, 4, POS, 1, 2, 7, 8, POS, 12, 12, 14, 14, POS, </s>	Analysis
Subtask	Target Sequence																					
<i>AE</i>	1, 2, 12, 12, </s>																					
<i>OE</i>	4, 4, 7, 8, 14, 14, </s>																					
<i>ALSC</i>	1, 2, POS, </s>																					
	12, 12, POS, </s>																					
<i>AOE</i>	1, 2, 4, 4, 7, 8, </s>																					
	12, 12, 14, 14, </s>																					
<i>AESC</i>	1, 2, POS, 12, 12, NEG, </s>																					
<i>Pair</i>	1, 2, 4, 4, 1, 2, 7, 8, 12, 12, 14, 14, </s>																					
<i>Triplet</i>	1, 2, 4, 4, POS, 1, 2, 7, 8, POS, 12, 12, 14, 14, POS, </s>																					
(M)LM	<div><div>The owners are great fun and the beer selection is worth staying for.</div><div>↓</div><div><div>Consistency prompt</div><div>The owners are great fun? [MASK] .</div></div><div><div>Polarity prompt</div><div>This is [MASK] .</div></div></div>	SentiPrompt: Sentiment Knowledge Enhanced Prompt-Tuning for Aspect-Based Sentiment Analysis																				

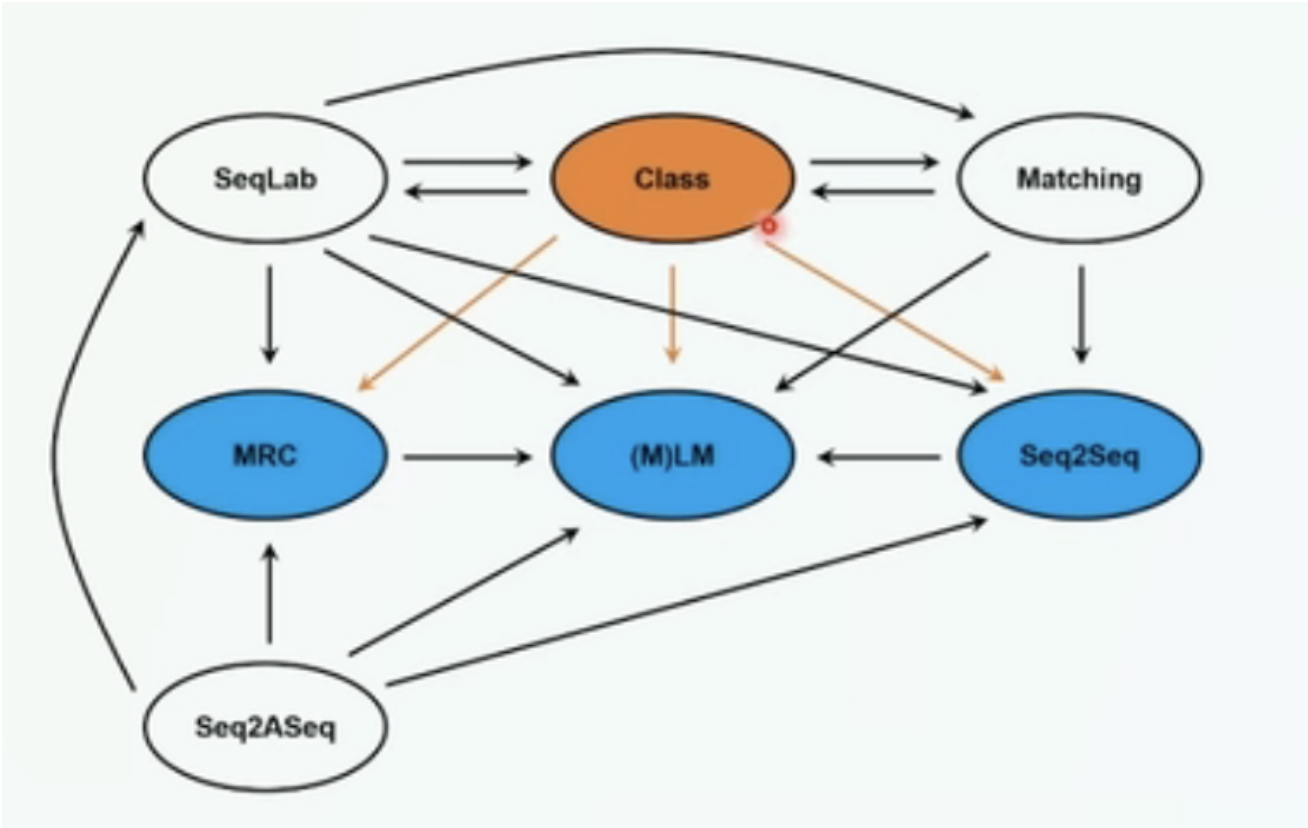
关系抽取

任务定义： 包括两部分， entity extraction, relation classification

SeqLab(entity extraction)

Class(relation classification)

范式迁移

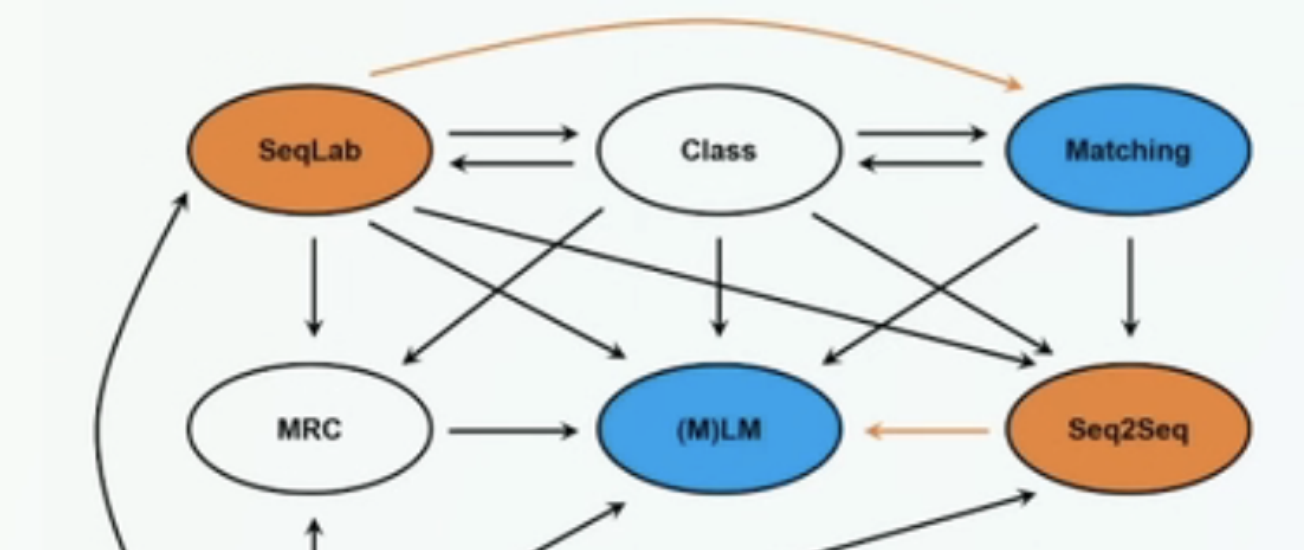


范式	图示	论文												
Seq2Seq		Extracting Relational Facts by an End-to-End Neural Model with Copy Mechanism												
MRC	<table border="1"> <tr> <td>Q1 Person:</td> <td>who is mentioned in the text?</td> <td>A: e_1</td> </tr> <tr> <td>Q2 Company:</td> <td>which companies did e_1 work for?</td> <td>A: e_2</td> </tr> <tr> <td>Q3 Position:</td> <td>what was e_1's position in e_2?</td> <td>A: e_3</td> </tr> <tr> <td>Q4 Time:</td> <td>During which period did e_1 work for e_2 as e_3?</td> <td>A: e_4</td> </tr> </table>	Q1 Person:	who is mentioned in the text?	A: e_1	Q2 Company:	which companies did e_1 work for?	A: e_2	Q3 Position:	what was e_1 's position in e_2 ?	A: e_3	Q4 Time:	During which period did e_1 work for e_2 as e_3 ?	A: e_4	Entity-Relation Extraction as Multi-turn Question Answering
Q1 Person:	who is mentioned in the text?	A: e_1												
Q2 Company:	which companies did e_1 work for?	A: e_2												
Q3 Position:	what was e_1 's position in e_2 ?	A: e_3												
Q4 Time:	During which period did e_1 work for e_2 as e_3 ?	A: e_4												
(M)LM	$T_{f_{e_s}}(x) = "x \text{ the [MASK] } e_s",$ $\mathcal{V}_{f_{e_s}} = \{ "person", "organization", \dots \}.$ $T_{f_{e_s, e_o}}(x) = "x \text{ } e_s \text{ [MASK] } e_o",$ $\mathcal{V}_{f_{e_s, e_o}} = \{ "s \text{ parent was", "was born in", \dots \}.$	PTR: Prompt Tuning with Rules for Text Classification												

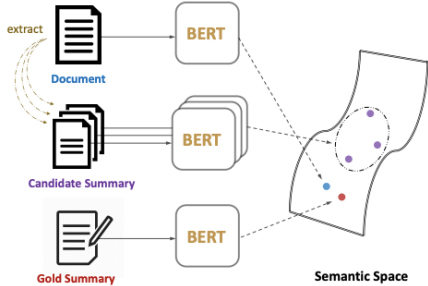
文本摘要

Extractive text summarization: 从文本中抽取一部分作为摘要 (SeqLab)

Abstractive text summarization: 根据文本自动生成(Seq2Seq)





范式	图示	论文
matching (extractive)	<div><p>Figure 1: MATCHSUM framework. We match the contextual representations of the document with gold summary and candidate summaries (extracted from the document). Intuitively, better candidate summaries should be semantically closer to the document, while the gold summary should be the closest.</p></div>	Extractive Summarization as Text Matching
(M)LM	<div><pre><!DOCTYPE html> <html> <title> <mask>12 </title> <body> ~ south korea on monday announced sweeping tax reforms , including income and corporate tax cuts to boost growth by stimulating sluggish private consumption and business investment . </body> </html></pre><p style="text-align: center;">↓</p><pre><!DOCTYPE html> <html></pre></div>	HTLM: Hyper-Text Pre-Training and Prompting of Language Models

```
<title> ~ South Korea Announces Tax Reforms To  
Boost Economic Growth ~ </title>  
<body>  
  ~ south korea on monday announced sweeping  
    tax reforms...  
</body>  
</html>
```

范式迁移趋势

<https://txsun1997.github.io/nlp-paradigm-shift/sankey.html>

传统： Class, SeqLab, Seq2ASeq

通用： Matching, MRC, Seq2Seq, (M)LM