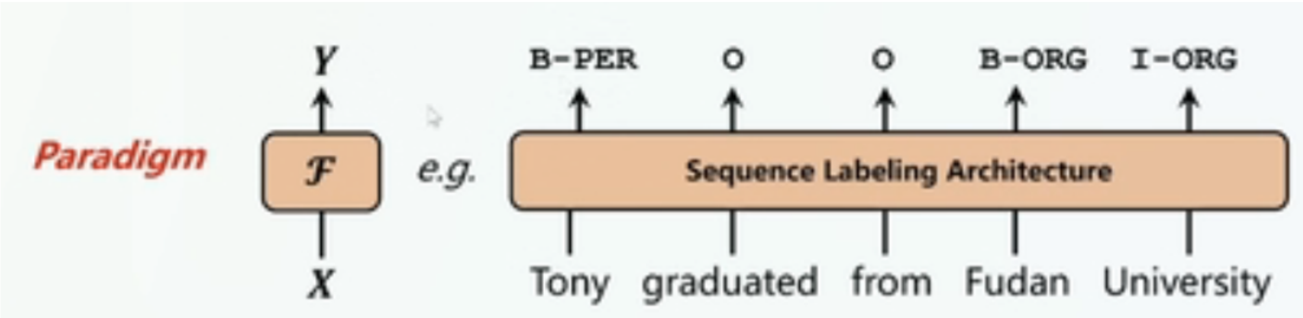
Paradigm Shift in Natural Language Processing

原文

范式

解决一类NLP任务的通用的框架



NLP七大范式

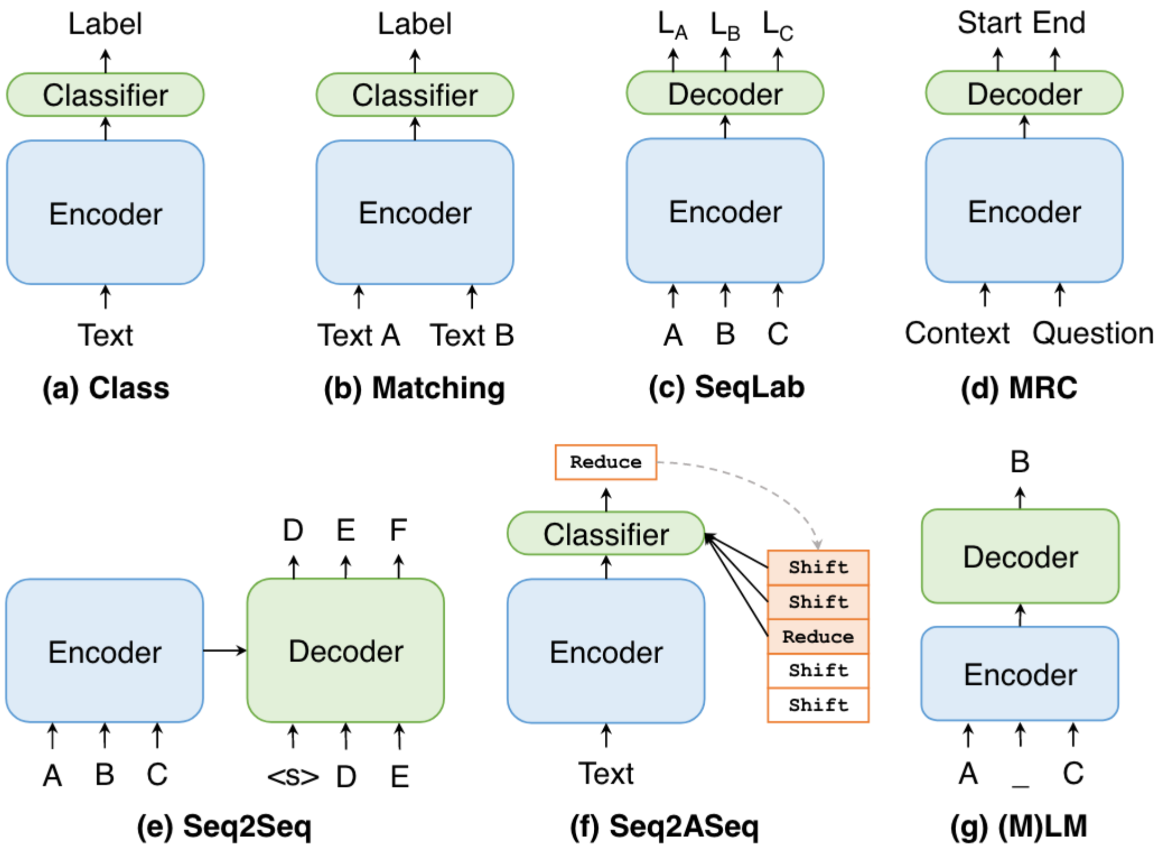
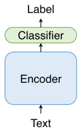
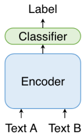
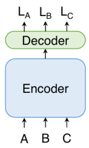
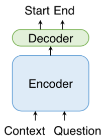
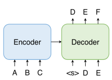
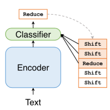
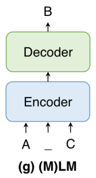


Figure 1: Illustration of the seven mainstream paradigms in NLP.

	图示		ENC	CLS/DEC	适用任务	备注
Class	 (a) Class	$\mathcal{Y} = \mathrm{CLS}(\mathrm{ENC}(\mathcal{X}))$	CNN/RNN/Transformer	(avg/max/attention) pooling + MLP	情感分析、主题分类、垃圾邮件检测....	
Matching	 (b) Matching	$\mathcal{Y} = \mathrm{CLS}(\mathrm{ENC}(\mathcal{X}_a, \mathcal{X}_b))$ Enc = enc2(enc1(x1), enc1(x2))	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)	 (c) SeqLab	$y_1, ..., y_n = \mathrm{DEC}(\mathrm{ENC}(x_1, ..., x_n))$	RNN/Transformer	CRF	NER, POS-Tagging	
MRC (machine reading comprehension)	 (d) MRC	$y_k, ..., y_{k+l} = \mathrm{DEC}(\mathrm{ENC}(\mathcal{X}_c, \mathcal{X}_q))$	CNN/RNN/Transformer	start/end position prediction	MRC	根据question，从context中抽取出连续的spans.
Seq2Seq	 (e) Seq2Seq	$y_1, ..., y_m = \mathrm{DEC}(\mathrm{ENC}(x_1, ..., x_n))$	CNN/RNN/Transformer	CNN/RNN/Transformer	机器翻译，对话	
Seq2ASeq (sequence-to-action-sequence)	 (f) Seq2ASeq	$\mathcal{A} = \mathrm{CLS}(\mathrm{ENC}(x), \mathcal{C})$	CNN/RNN/Transformer	基于输入文本和当前栈的状态预测action	Dependency parsing	CLS除了有encoder的输出，还要维护一个栈的状态

(M)LM		$\mathrm{LM}: x_k = \mathrm{DEC}(x_1, ..., x_{k-1})$ $\mathrm{MLM}: \bar{x} = \mathrm{DEC}(\mathrm{ENC}(\tilde{x}))$	CNN/RNN/Transformer	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

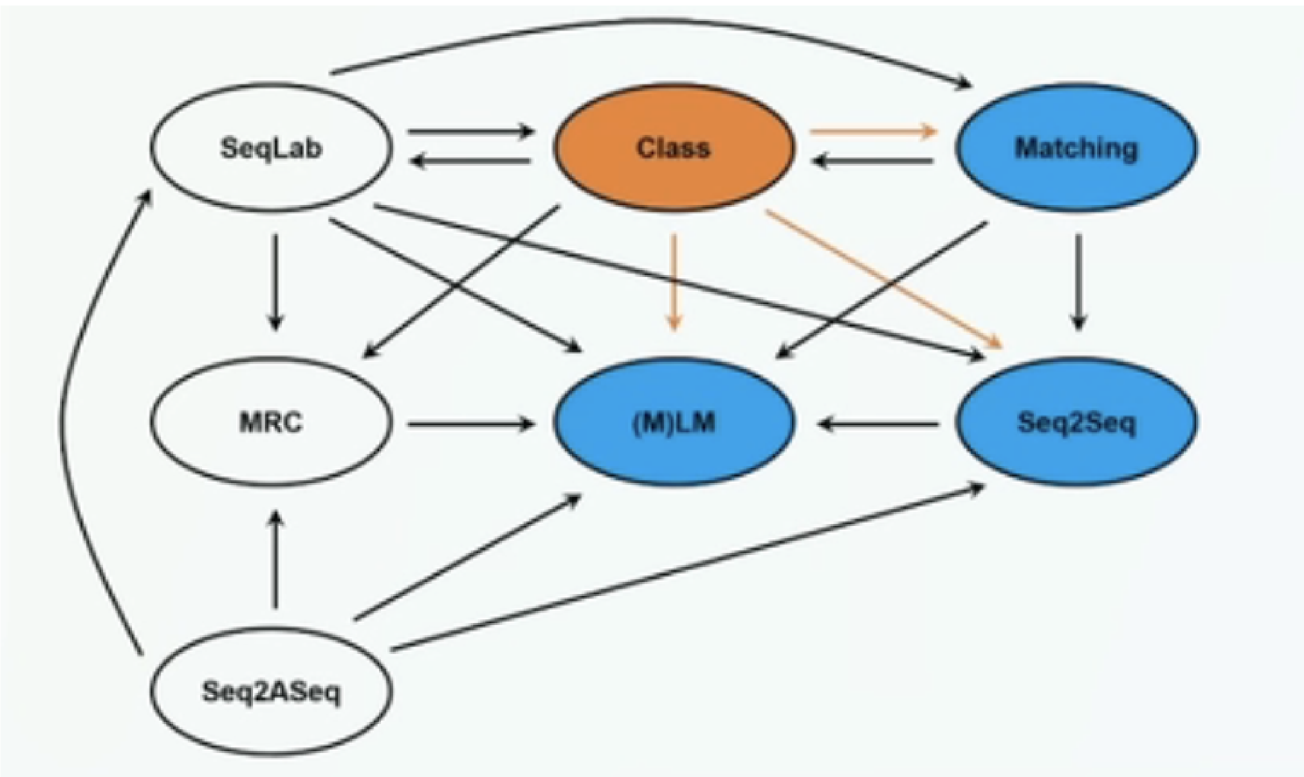
matching找到相关的document

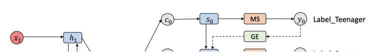
MRC从document中抽取出answer

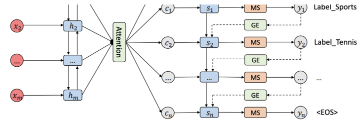
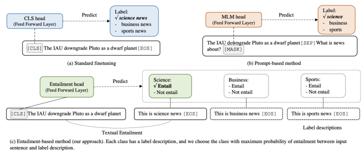
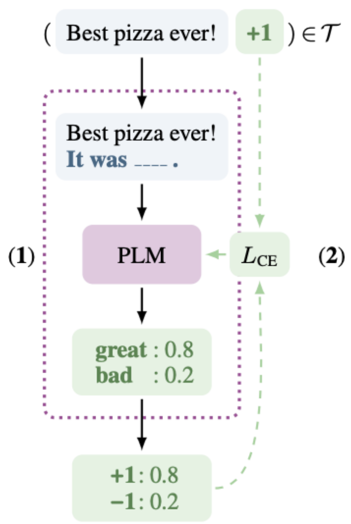
范式迁移

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

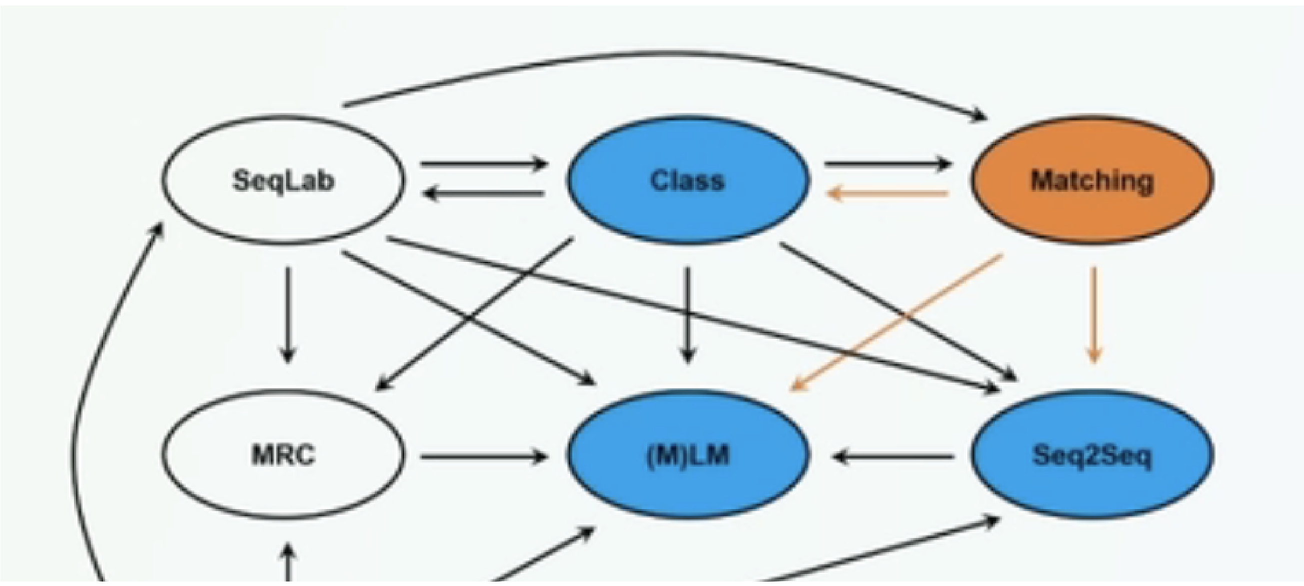


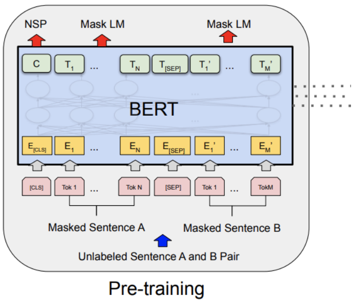
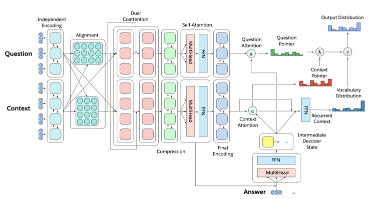
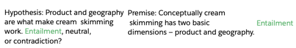
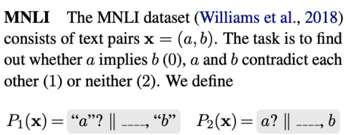
范式	图示	备注	论文
Seq2Se		多标签分类任务中，Class	SGM: sequence generation model

q		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.	for multi-label classification. 2018 coling best paper
Matching		用原始句子和标签描述做一个NLI分类（entail/not entail）	Entailment as few-shot learner.
(M)LM			Exploiting cloze-questions for few-shot text classification and natural language inference.

NLI

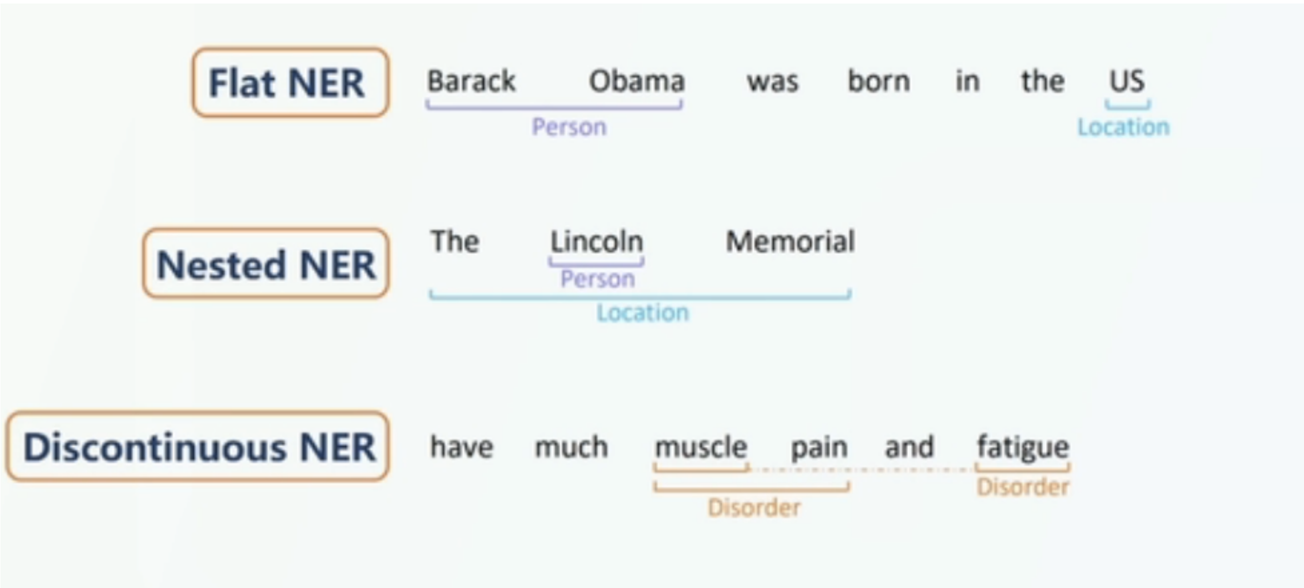
自然语言推断任务，即给出一对句子, 判断两个句子是entailement/contradiction/neutral.

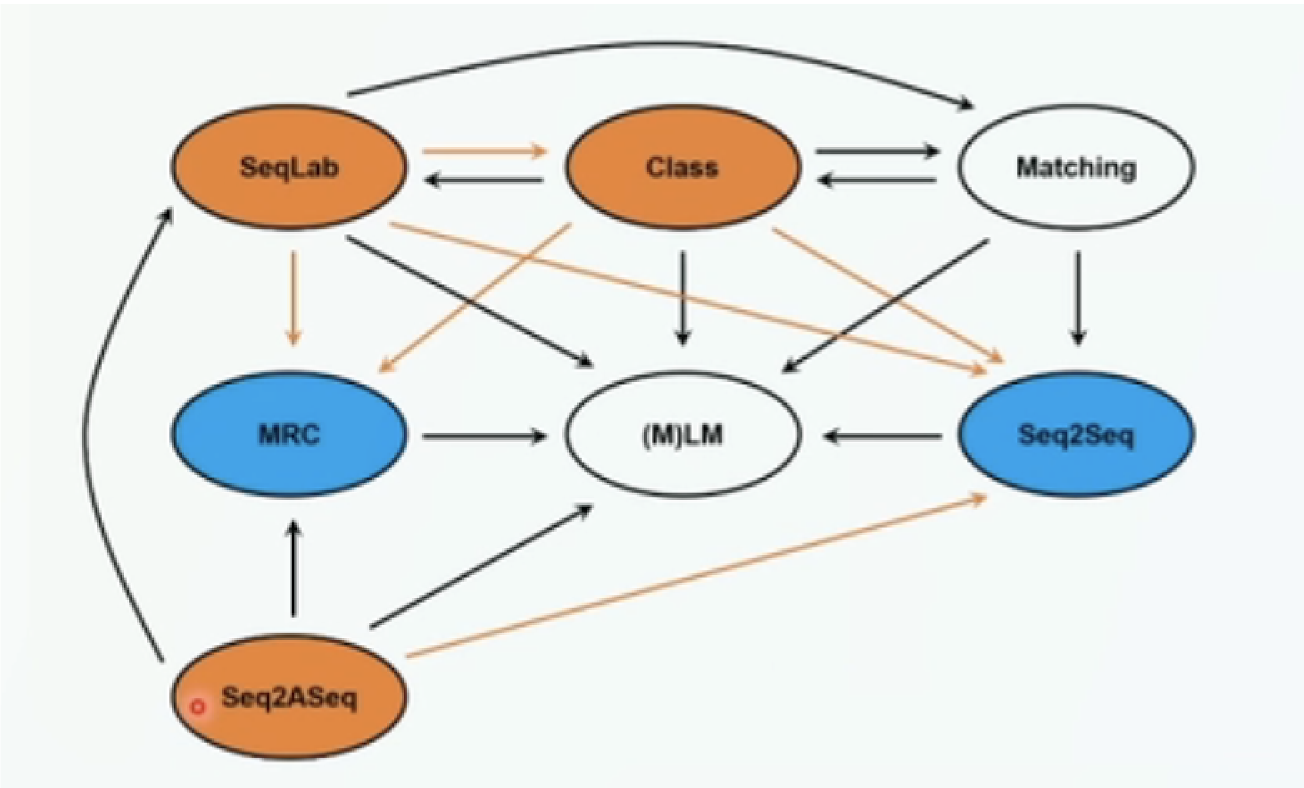


范式	图示	备注	论文
Class			
Seq2Seq		 把所有任务都形式化为QA的形式，context, question,answer	The Natural Language Decathlon: Multitask Learning as Question Answering
(M)LM	 **MNLI** The MNLI dataset (Williams et al., 2018) consists of text pairs $\mathbf{x} = (a, b)$. The task is to find out whether a implies b (0), a and b contradict each other (1) or neither (2). We define $P_1(\mathbf{x}) = $ "a"? ‖ ____, "b" $P_2(\mathbf{x}) = a?$ ‖ ____, 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)

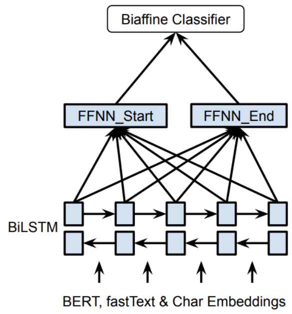
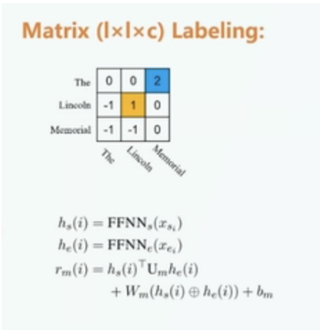
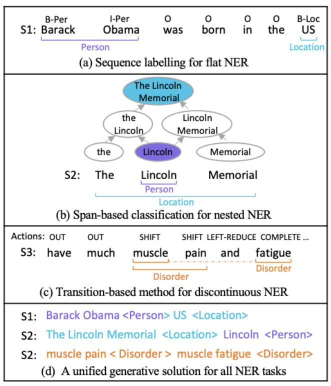
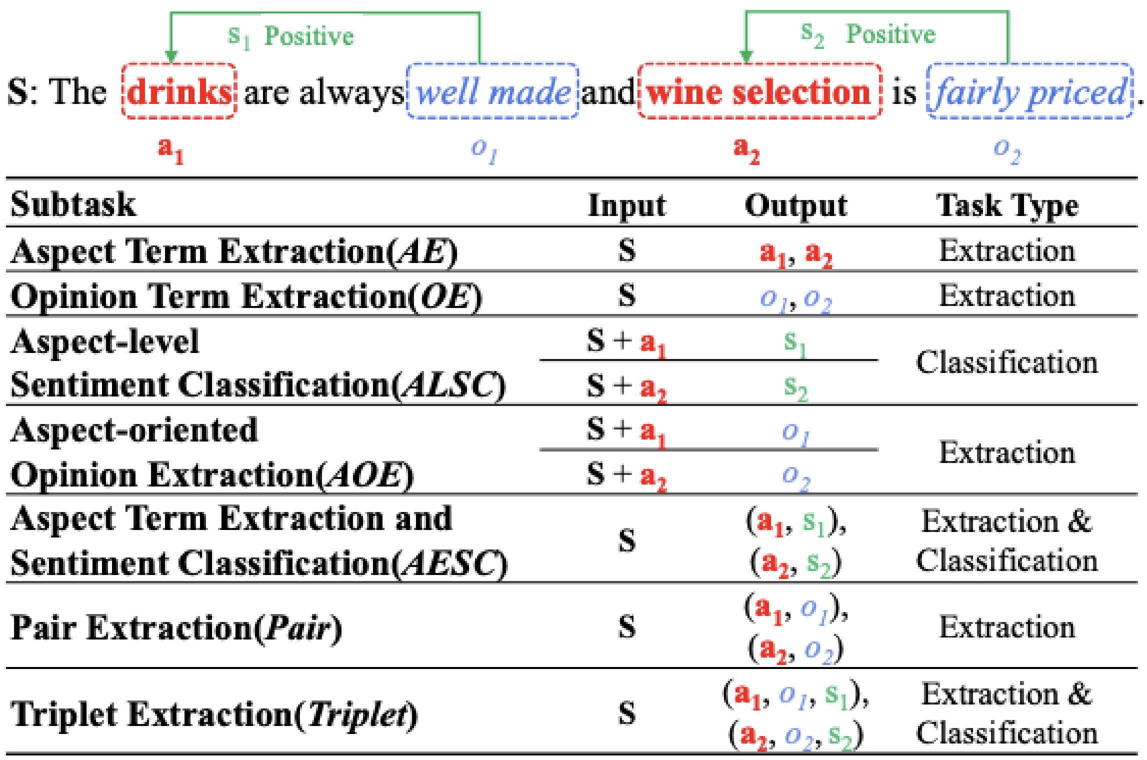
<table>
<tr><th>范式</th><th>图示</th><th>备注</th><th>论文</th></tr>
<tr><td>Class(flat & nested)</td><td>

Figure 1: The network architectures of our system.

</td><td>Biaffine: https://zhuanlan.zhihu.com/p/369851456</td><td>Named Entity Recognition as Dependency Parsing</td></tr>
<tr><td>MRC(flat & Nested)</td><td><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>
Table 1: Examples for transforming different entity categories to question queries.</td><td></td><td>A unified MRC framework for named entity recognition</td></tr>
<tr><td>Seq2Seq(All)</td><td>
</td><td>BART+pointer network</td><td>A unified generative framework for various NER subtasks.</td></tr>
</table>

ABSA

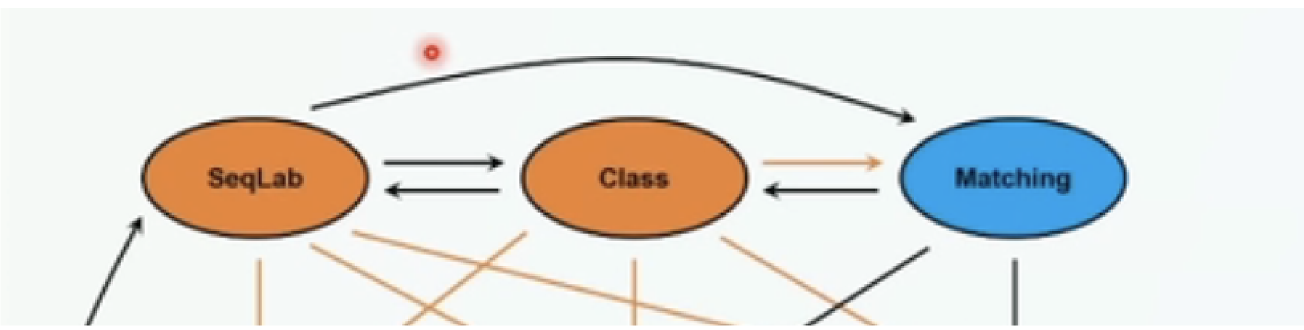
任务定义：

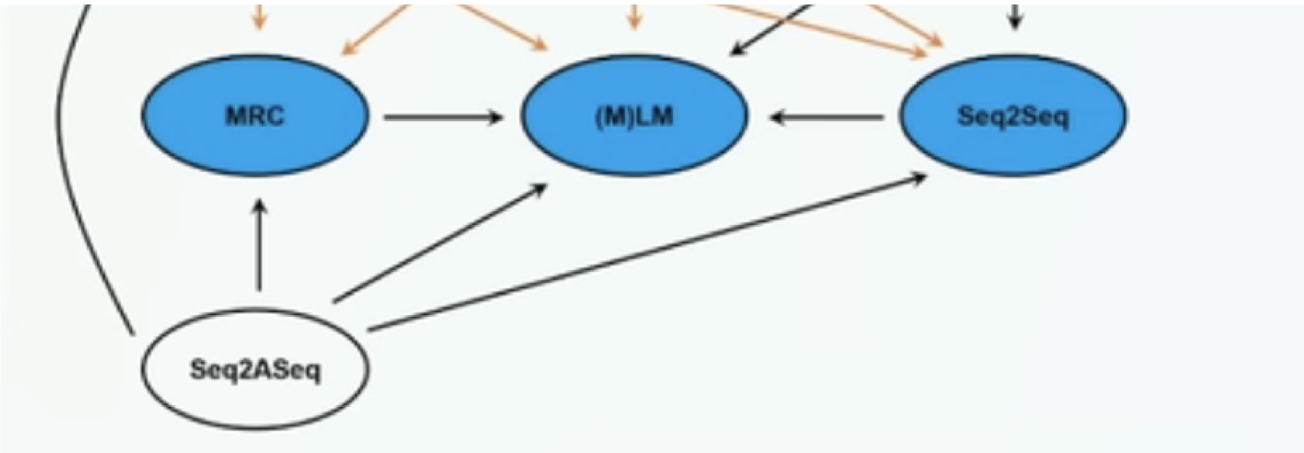


Subtask	Input	Output	Task Type
Aspect Term Extraction(*AE*)	S	a_1, a_2	Extraction
Opinion Term Extraction(*OE*)	S	o_1, o_2	Extraction
Aspect-level Sentiment Classification(*ALSC*)	S + a_1	s_1	Classification
	S + a_2	s_2	
Aspect-oriented Opinion Extraction(*AOE*)	S + a_1	o_1	Extraction
	S + a_2	o_2	
Aspect Term Extraction and Sentiment Classification(*AESC*)	S	(a_1, s_1), (a_2, s_2)	Extraction & Classification
Pair Extraction(*Pair*)	S	(a_1, o_1), (a_2, o_2)	Extraction
Triplet Extraction(*Triplet*)	S	(a_1, o_1, s_1), (a_2, o_2, s_2)	Extraction & Classification

Figure 1: Illustration of seven ABSA subtasks.

范式迁移：

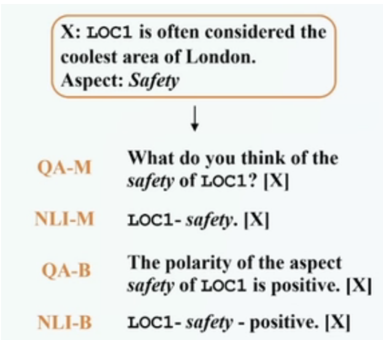
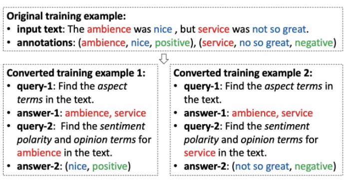
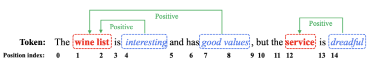


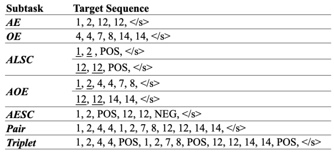


传统方法

SeqLab(AE, OE, AOE)

Class(ALSC)

范式	图示	备注	论文
Matching(ALSC)			Utilizing BERT for Aspect-Based Sentiment Analysis via Constructing Auxiliary Sentence
MRC(All)	 Figure 3: Dataset conversion.		A joint training dual-mrc framework for aspect based sentiment analysis.
Seq2Seq(All)		BART + pointer network	A Unified Generative Framework for Aspect-Based Sentiment

<table>
<tr>
<td></td>
<td>
<table>
<tr><th>Subtask</th><th>Target Sequence</th></tr>
<tr><td>AE</td><td>1, 2, 12, 12, </s></td></tr>
<tr><td>OE</td><td>4, 4, 7, 8, 14, 14, </s></td></tr>
<tr><td rowspan="2">ALSC</td><td>1, 2 , POS, </s></td></tr>
<tr><td>12, 12, POS, </s></td></tr>
<tr><td rowspan="2">AOE</td><td>1, 2, 4, 4, 7, 8, </s></td></tr>
<tr><td>12, 12, 14, 14, </s></td></tr>
<tr><td>AESC</td><td>1, 2, POS, 12, 12, NEG, </s></td></tr>
<tr><td>Pair</td><td>1, 2, 4, 4, 1, 2, 7, 8, 12, 12, 14, 14, </s></td></tr>
<tr><td>Triplet</td><td>1, 2, 4, 4, POS, 1, 2, 7, 8, POS, 12, 12, 14, 14, POS, </s></td></tr>
</table>
</td>
<td></td>
<td>Analysis</td>
</tr>
<tr>
<td>(M)LM</td>
<td>

</td>
<td></td>
<td>SentiPrompt: Sentiment Knowledge Enhanced Prompt-Tuning for Aspect-Based Sentiment Analysis</td>
</tr>
</table>

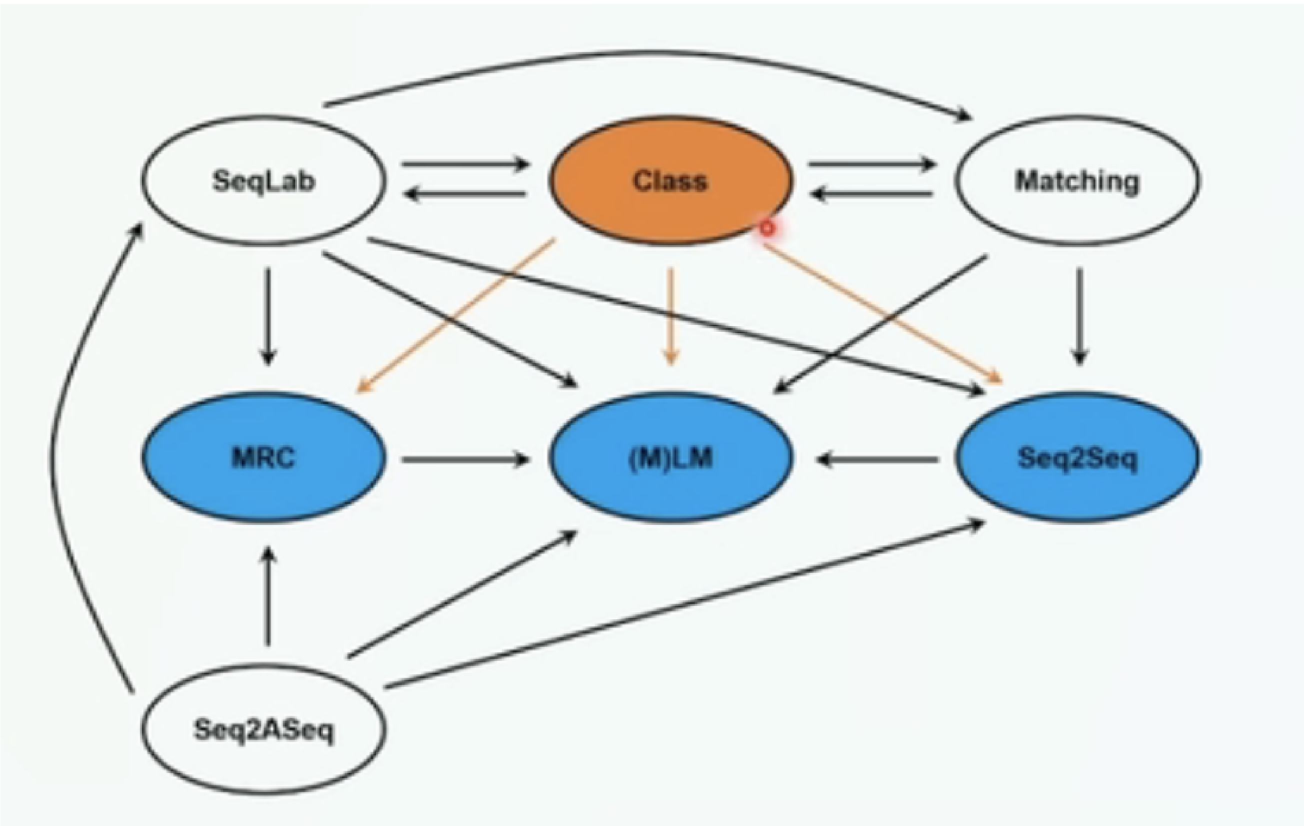
关系抽取

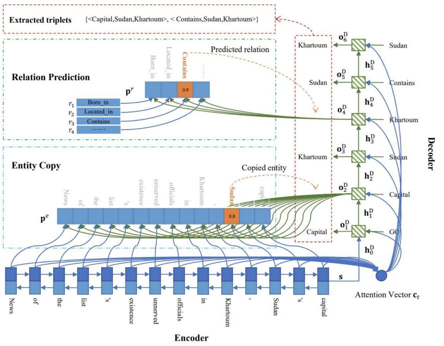
任务定义： 包括两部分，entity extraction, relation classification

SeqLab(entity extraction)

Class(relation classification)

范式迁移

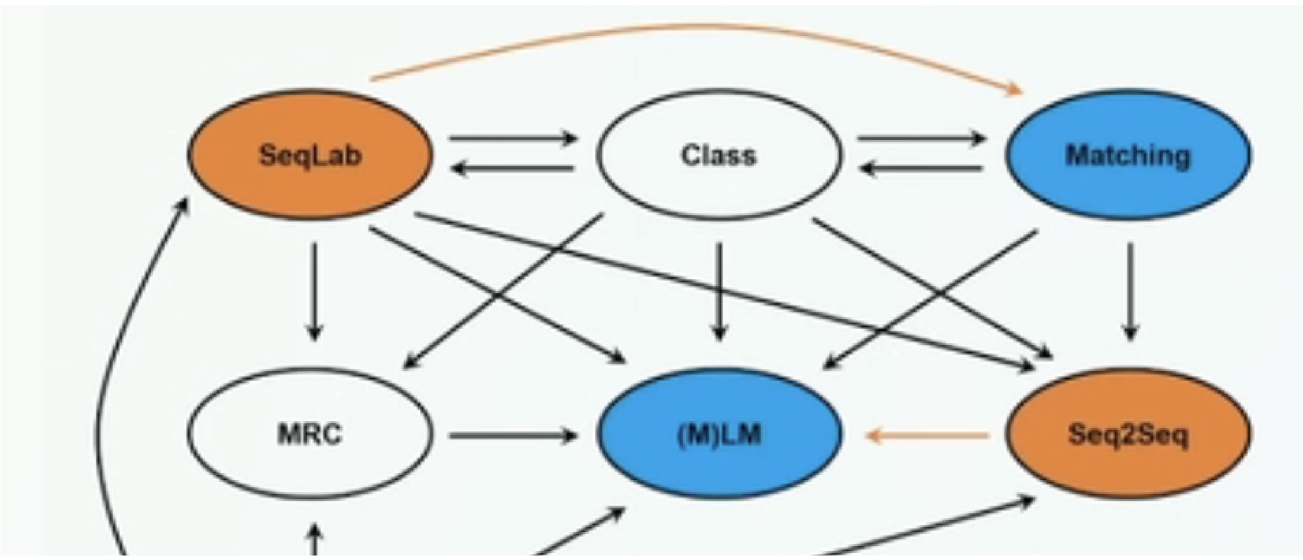


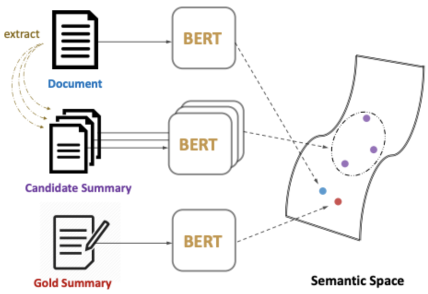
范式	图示	论文
Seq2Seq		Extracting Relational Facts by an End-to-End Neural Model with Copy Mechanism
MRC	Q1 Person: who is mentioned in the text? A: e_1 Q2 Company: which companies did e_1 work for? A: e_2 Q3 Position: what was e_1's position in e_2? A: e_3 Q4 Time: During which period did e_1 work for e_2 as e_3 A: e_4	Entity-Relation Extraction as Multi-turn Question Answering
(M)LM	$T_{f_{e_s}}(x) = \text{“}x \text{ the } \texttt{[MASK]}\ e_s\text{”}, \quad \mathcal{V}_{f_{e_s}} = \{\text{“person”}, \text{“organization”}, \ldots\}.$ $T_{f_{e_s,e_o}}(x) = \text{“ } x\ e_s\ \texttt{[MASK]}\ e_o\text{”}, \quad \mathcal{V}_{f_{e_s,e_o}} = \{\text{“'s parent was”}, \text{“was born in”}, \ldots\}.$	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)	 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.	Extractive Summarization as Text Matching
(M)LM	<!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> ↓ <!DOCTYPE html> <html>	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