DataFunSummit

自然语言处理

在线峰会

产业创新与实践 论坛

2021.07.10（周六） 09：00~17：30

2021

your logo | DataFunSummit

K12智能批改技术的探索与实践

李超 腾讯教育高级研究员

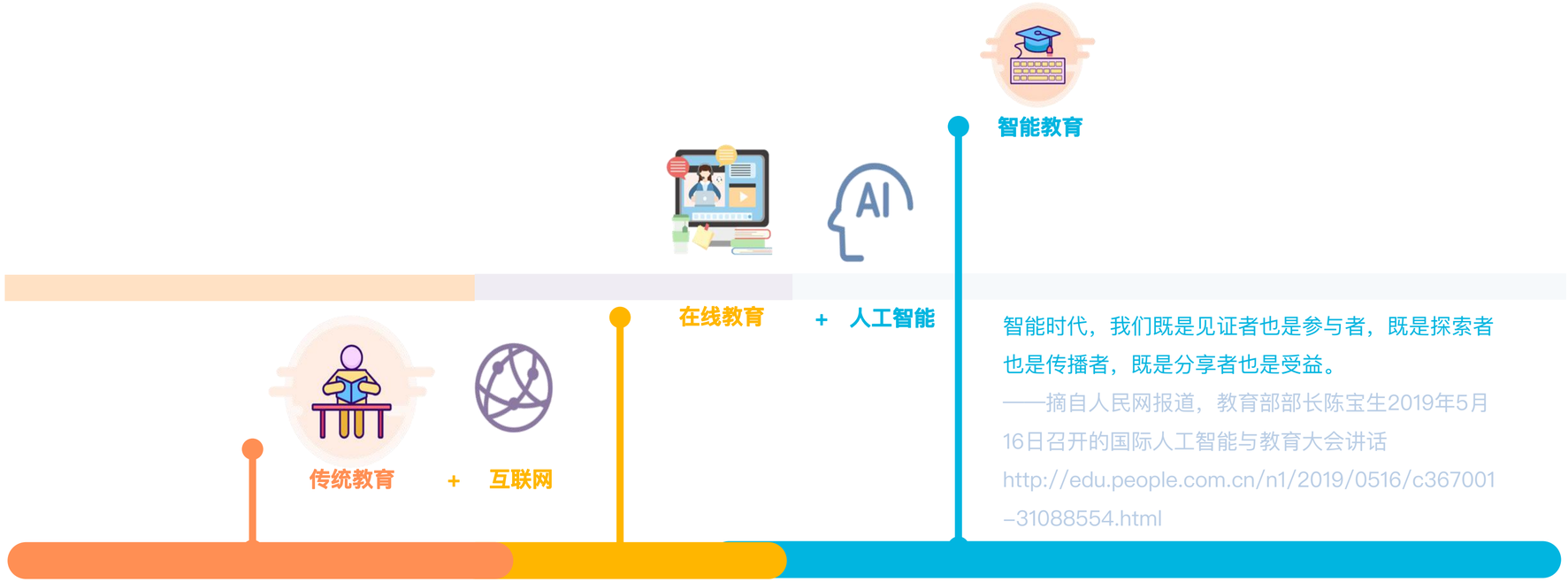
目录

CONTENTS

01
背景介绍

Subject

背景介绍

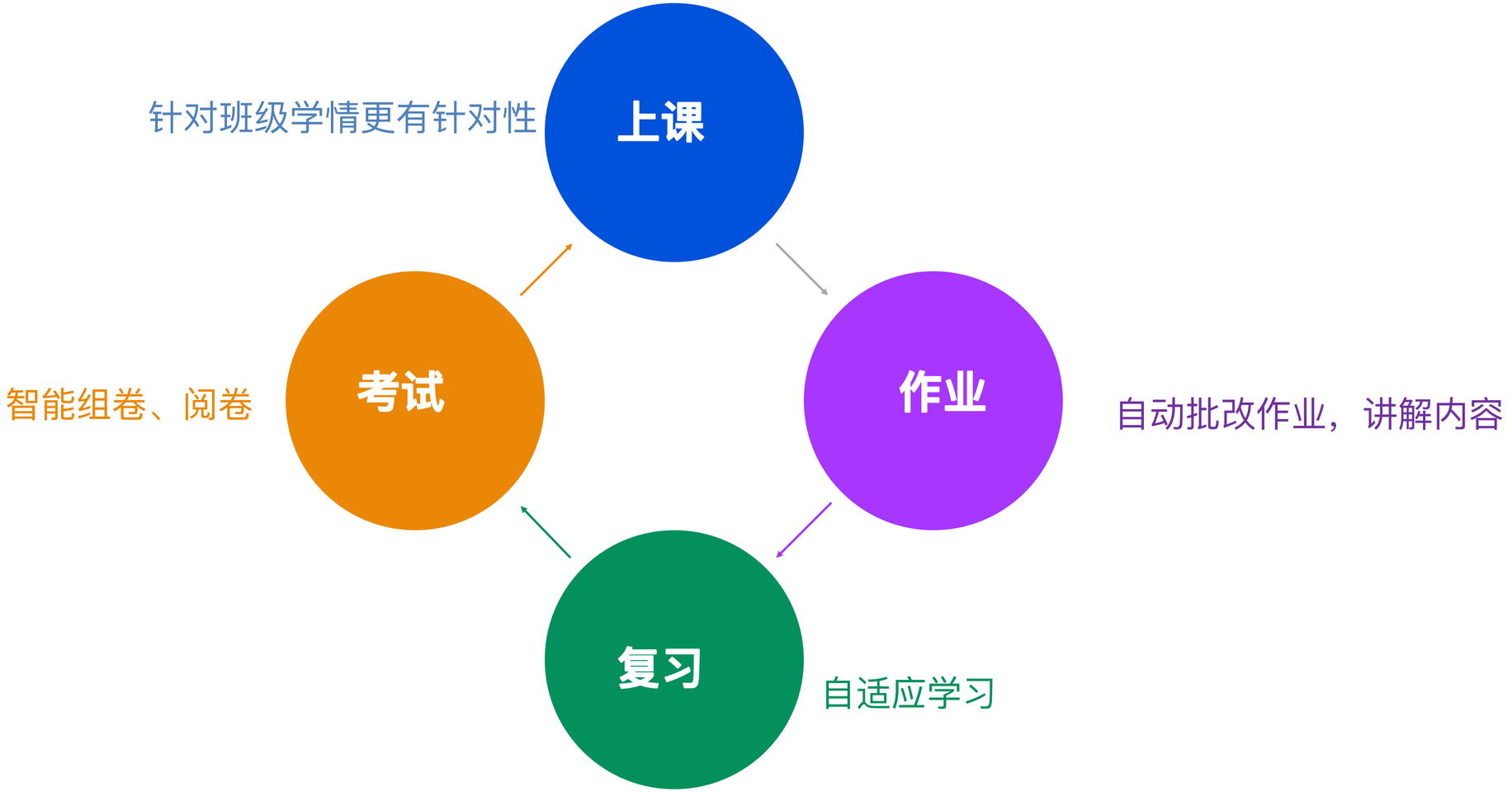


智能时代，我们既是见证者也是参与者，既是探索者也是传播者，既是分享者也是受益。

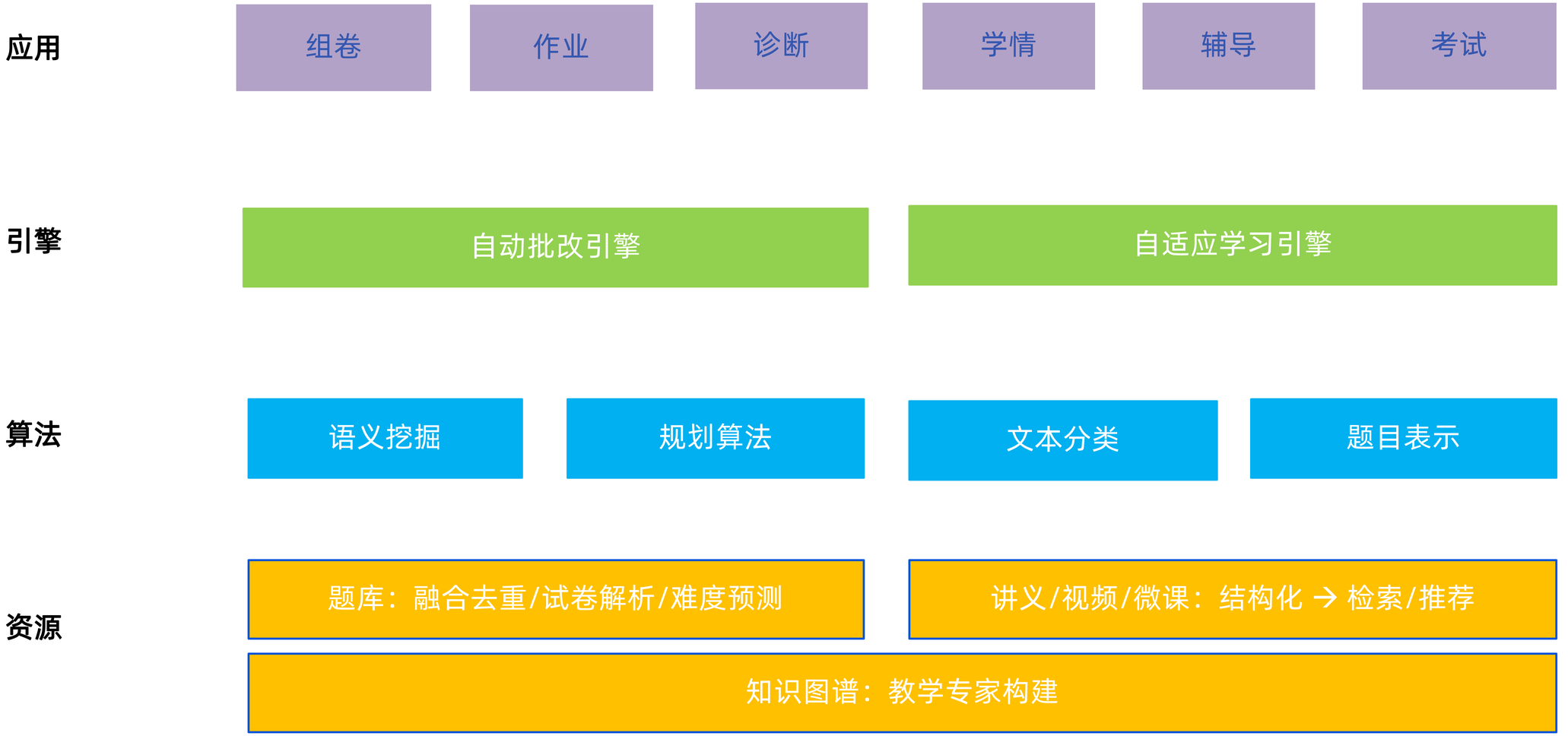
——摘自人民网报道，教育部部长陈宝生2019年5月16日召开的国际人工智能与教育大会讲话

http://edu.people.com.cn/n1/2019/0516/c367001-31088554.html

背景介绍



背景介绍



02
典型问题

Subject

背景介绍-教育图谱

知识点体系

主要依据课标+教材编写顺序

题模体系

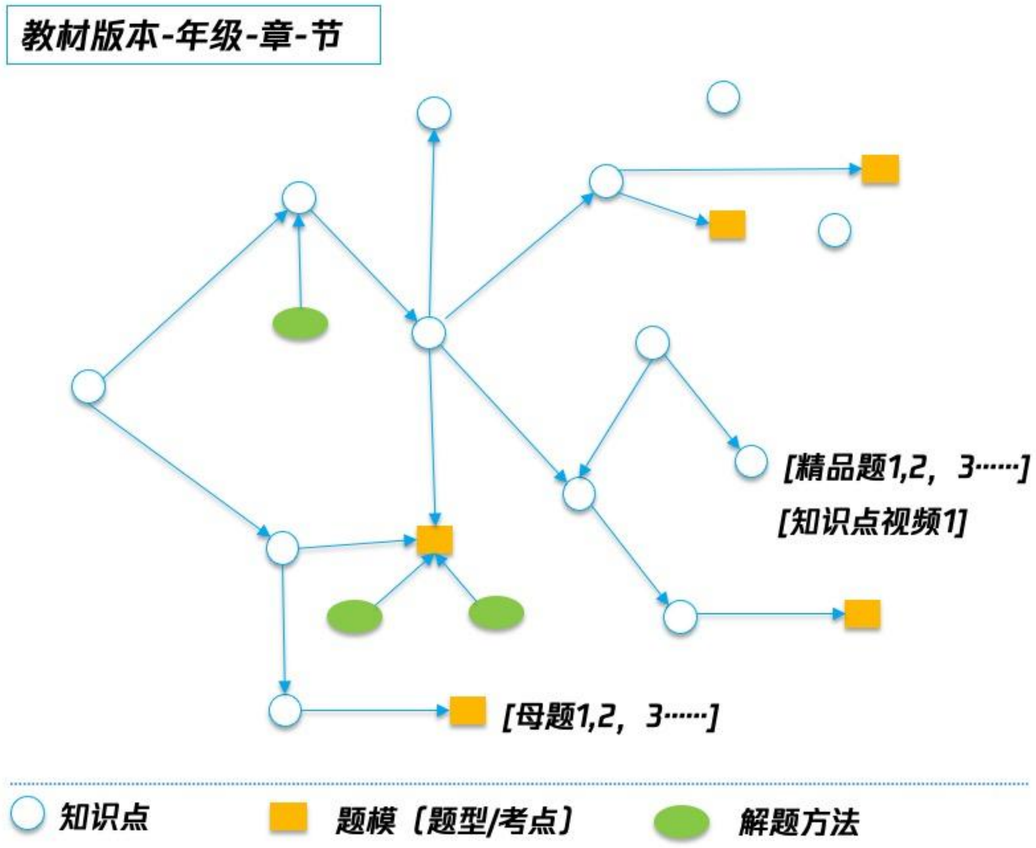
面向考试&解题，提炼关键题型、考法、解题方法

试卷解析

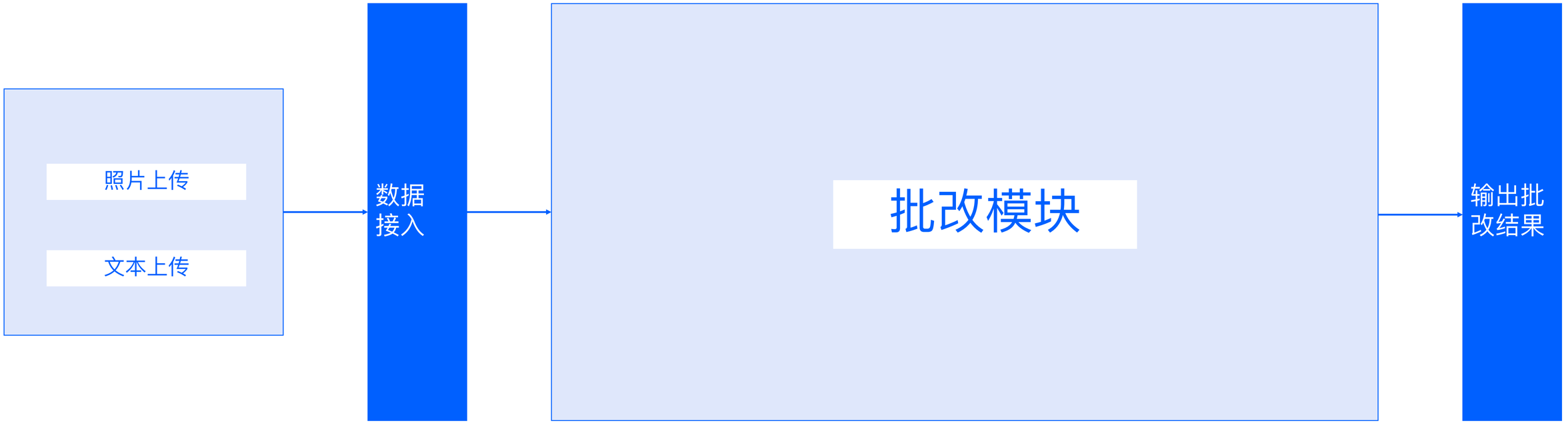
题目知识点识别

题目难度识别

视频知识点识别



背景介绍-批改

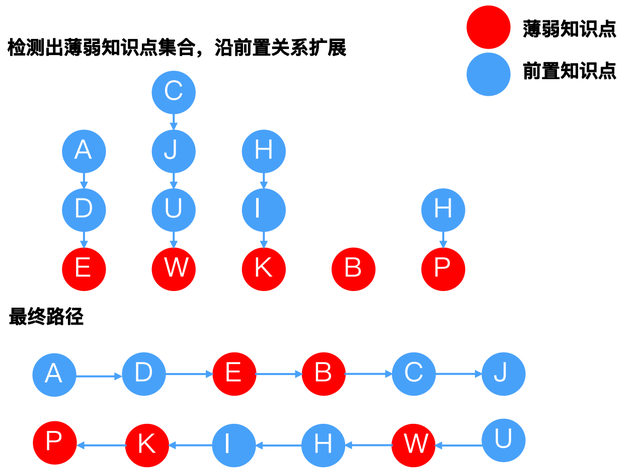


背景介绍

学情画像
　　知识点掌握概率

学习路径规划

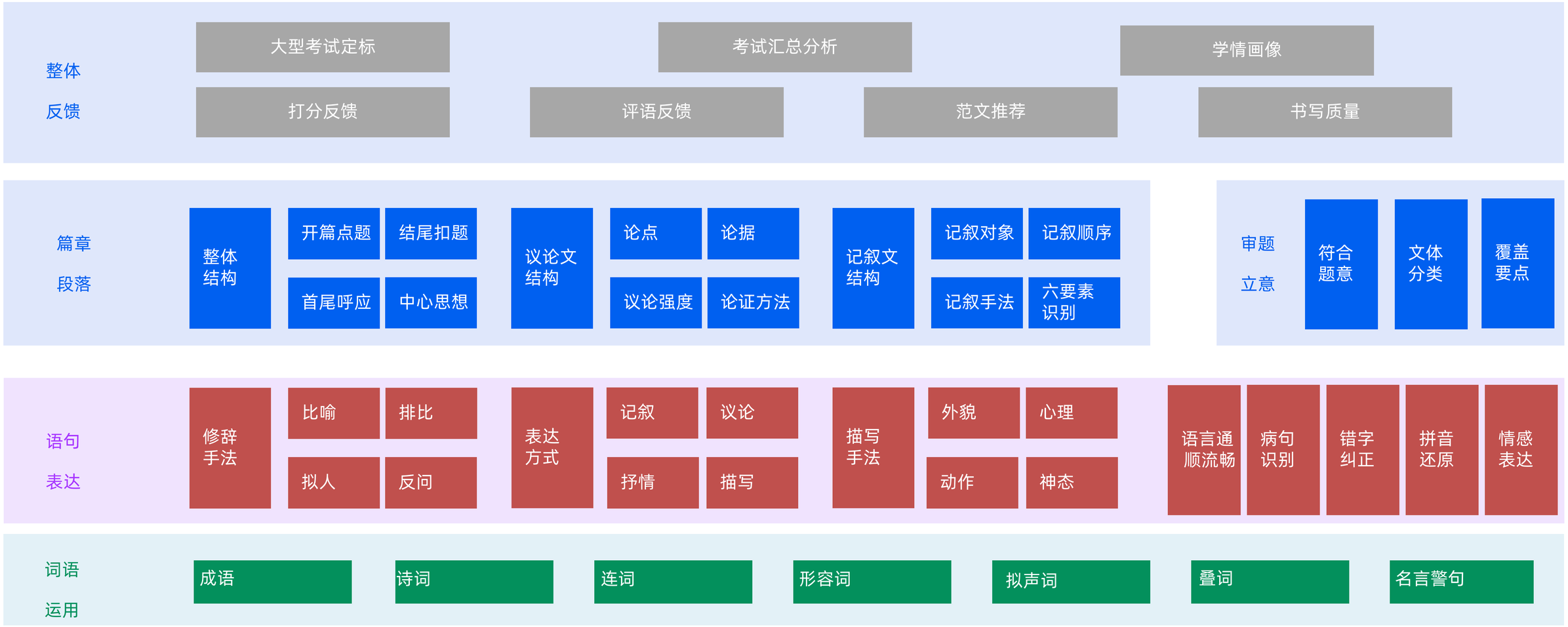
相似重复题目推荐



03
技术方案

Subject

技术方案-作文批改



技术方案-拼写纠错

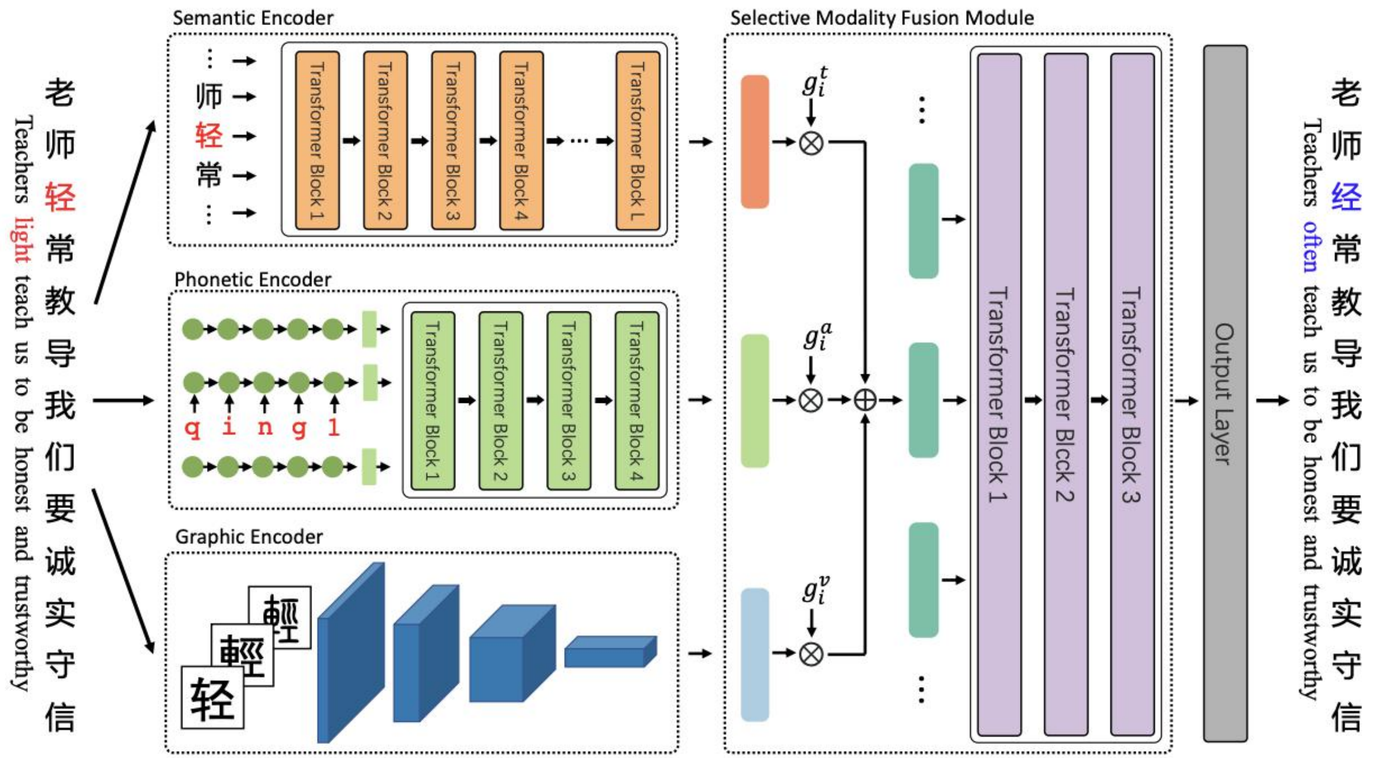


Figure 1: Architecture overview of the REALISE model. The semantic, phonetic and graphic encoders, are used to capture the information in textual, acoustic and visual modalities. The fusion module selectively fuses the information from three encoders. In the example input, to correct the erroneous character, "轻" (qīng, light), we need not only the contextual text information, but also the phonetic and graphic information of the character itself.

技术方案-拼写纠错

Dataset	Method	Detection Level				Correction Level			
		Acc	Pre	Rec	F1	Acc	Pre	Rec	F1
SIGHAN13	Sequence Labeling (Wang et al., 2018)	-	54.0	69.3	60.7	-	-	-	52.1
	FASpell (Hong et al., 2019)	63.1	76.2	63.2	69.1	60.5	73.1	60.5	66.2
	BERT (Cheng et al., 2020)	-	79.0	72.8	75.8	-	77.7	71.6	74.6
	SpellGCN (Cheng et al., 2020)	-	80.1	74.4	77.2	-	78.3	72.7	75.4
	SpellGCN† (Our reimplementation)	78.8	85.7	78.8	82.1	77.8	84.6	77.8	81.0
	BERT†	77.0	85.0	77.0	80.8	75.2	83.0	75.2	78.9
	REALISE†	**82.7**	**88.6**	**82.5**	**85.4**	**81.4**	**87.2**	**81.2**	**84.1**
SIGHAN14	Sequence Labeling (Wang et al., 2018)	-	51.9	66.2	58.2	-	-	-	56.1
	FASpell (Hong et al., 2019)	70.0	61.0	53.5	57.0	69.3	59.4	52.0	55.4
	SpellGCN (Cheng et al., 2020)	-	65.1	69.5	67.2	-	63.1	67.2	65.3
	BERT	75.7	64.5	68.6	66.5	74.6	62.4	66.3	64.3
	REALISE	**78.4**	**67.8**	**71.5**	**69.6**	**77.7**	**66.3**	**70.0**	**68.1**
SIGHAN15	KUAS (Chang et al., 2015)	53.2	57.5	24.6	34.4	51.5	53.7	21.1	30.3
	NTOU (Chu and Lin, 2015)	42.2	42.2	41.8	42.0	39.0	38.1	35.2	36.6
	NCTU-NTUT (Wang and Liao, 2015)	60.1	71.7	33.6	45.7	56.4	66.3	26.1	37.5
	HanSpeller++ (Zhang et al., 2015)	70.1	80.3	53.3	64.0	69.2	79.7	51.5	62.5
	LMC (Xie et al., 2015)	54.6	63.8	21.5	32.1	52.3	57.9	16.7	26.0
	Sequence Labeling (Wang et al., 2018)	-	56.6	69.4	62.3	-	-	-	57.1
	FASpell (Hong et al., 2019)	74.2	67.6	60.0	63.5	73.7	66.6	59.1	62.6
	Soft-Masked BERT (Zhang et al., 2020)	80.9	73.7	73.2	73.5	77.4	66.7	66.2	66.4
	SpellGCN (Cheng et al., 2020)	-	74.8	80.7	77.7	-	72.1	77.7	75.9
	BERT	82.4	74.2	78.0	76.1	81.0	71.6	75.3	73.4
	REALISE	**84.7**	**77.3**	**81.3**	**79.3**	**84.0**	**75.9**	**79.9**	**77.8**

Table 3: The performance of our model and all baseline models on SIGHAN test sets. The "†" symbol means we apply post-processing (Section 4.2) to the model outputs on SIGHAN13. Results of REALISE on all SIGHAN test sets outperforms all the corresponding baselines with a significance level $p < 0.05$.

技术方案-病句识别

序列标注的方法 seq2label

Grammarly: Grammatical Error Correction: Tag, Not Rewrite

标签类型：保持不变（$KEEP）,删除（ $DELETE）,增加（ $APPEND_t）,替换（ $REPLACE_t），其中增加和替换的t为指定的单词

A	$KEEP	
ten	$MERGE_HYPHEN	
years	$NOUN_NUMBER_SINGULAR	
old	$KEEP	A ten-year-old boy goes to school.
boy	$KEEP	
go	$VERB_FORM_VB_VBZ	
to	$KEEP	
school	$APPEND_{.}	

其他：类机器翻译的方法 seq2seq

技术方案-病句识别

Results (span-level correction)											
#	User	Entries	Date of Last Entry	Team Name	F_0.5 ▲	P ▲	R ▲	TP ▲	FP ▲	FN ▲	Detailed Results
1	JackieHan	70	04/30/21	Tencent Cloud Xiaowei	77.21 (1)	84.87 (5)	56.73 (13)	2996.0 (15)	534.0 (26)	2285.0 (24)	View
2	linhw	10	12/22/20	USTC-linhw	76.75 (2)	84.97 (4)	55.35 (17)	2922.0 (17)	517.0 (28)	2357.0 (18)	View
3	zxlxdf	84	01/26/21	XDF-NLP	76.59 (3)	85.86 (1)	53.48 (21)	2733.0 (22)	450.0 (32)	2377.0 (16)	View
4	WangLicheng	53	01/15/21	WLC	76.59 (3)	84.45 (8)	55.81 (16)	2906.0 (19)	535.0 (25)	2301.0 (22)	View
5	thx	16	01/10/21	Tencent	75.71 (4)	82.85 (10)	56.29 (15)	3025.0 (14)	626.0 (23)	2349.0 (20)	View
6	zhoujunyu9102	108	02/09/21	ZJY	75.59 (5)	85.71 (2)	51.34 (25)	2735.0 (21)	456.0 (31)	2592.0 (7)	View
7	Lenovo_coretech	2	12/30/20	Smart EDU-Lenovo	75.39 (6)	84.78 (6)	52.23 (22)	2724.0 (23)	489.0 (29)	2491.0 (11)	View
8	TAL_ML_team	9	12/22/20	TAL ML team	75.17 (7)	85.70 (3)	50.39 (27)	2608.0 (26)	435.0 (33)	2568.0 (9)	View
9	zaid	27	12/18/20	CVTE-NLP	75.14 (8)	84.75 (7)	51.69 (23)	2662.0 (24)	479.0 (30)	2488.0 (12)	View
10	pdzong	23	07/15/20	becorrect.com	73.90 (9)	83.02 (9)	51.35 (24)	2581.0 (28)	528.0 (27)	2445.0 (14)	View
11	valuable	11	11/18/20	leo	73.67 (10)	78.81 (13)	58.42 (9)	3265.0 (8)	878.0 (18)	2324.0 (21)	View
12	kyriez	90	12/03/20	Mobvoi NLP	72.78 (11)	82.07 (11)	50.10 (28)	2599.0 (27)	568.0 (24)	2589.0 (8)	View

技术方案-内容理解

修辞识别

比喻识别：本体/喻体/比喻词

排比识别：句对分类

表达方式识别

描写方式识别

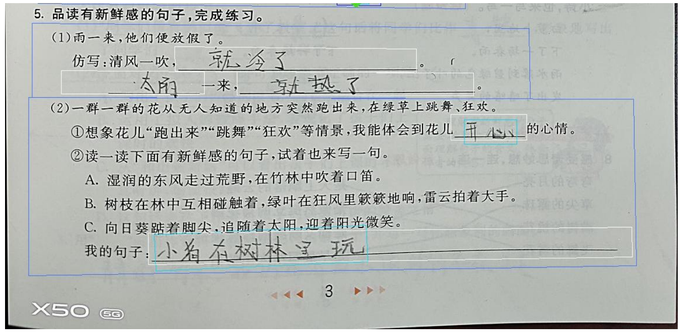
记叙文六要素识别

论证方式识别

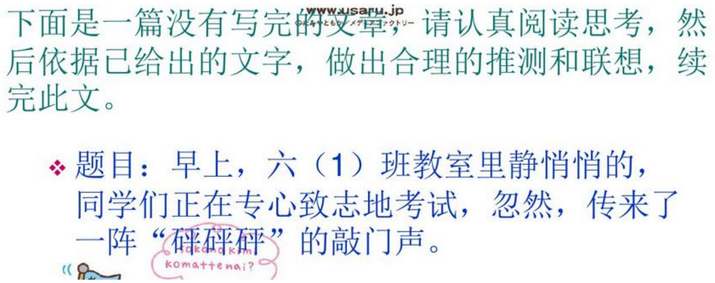
议论强度识别

技术方案-其他写作

仿写



续写



看图作文

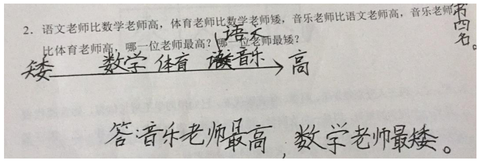
句式判断模型

连续性判断模型

图像内容识别

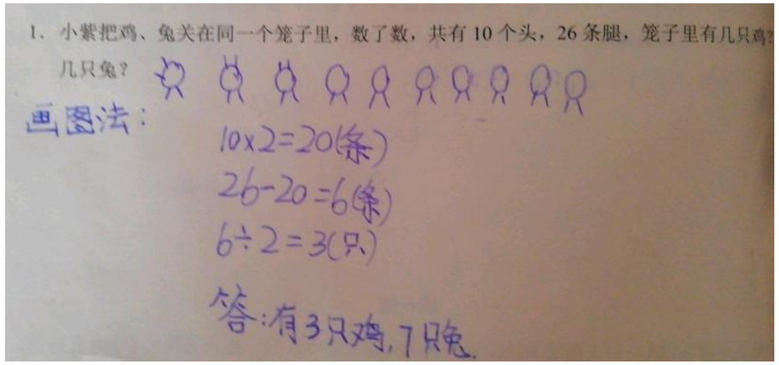
技术方案-应用题批改

文本答案判断



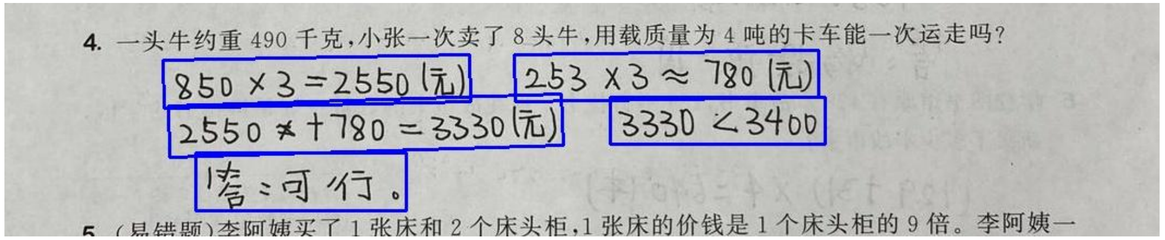
语义一致性模型

数值答案批改



数值答案批改

算式切题判定



自动解题模型

技术方案-填空题批改

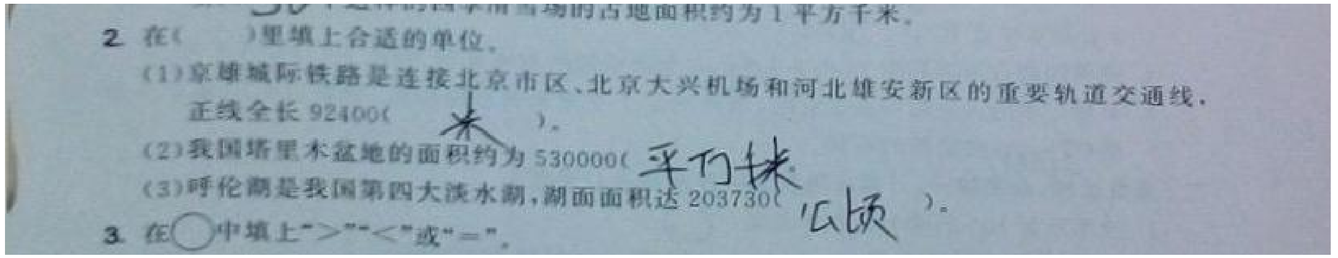
问题

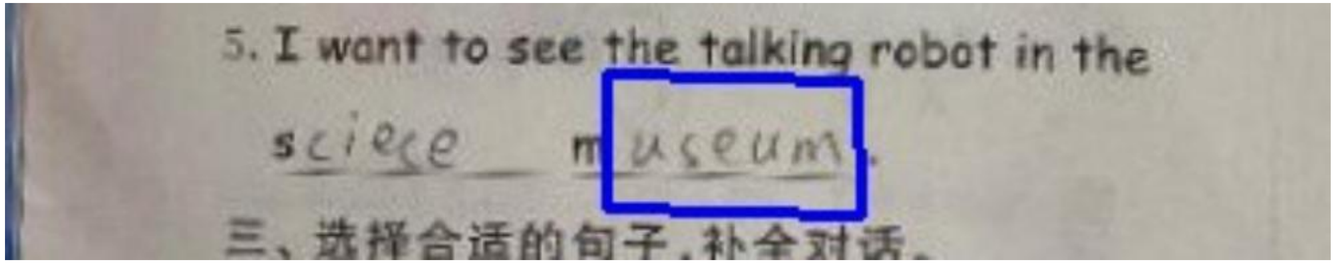
OCR错误

方案

文本纠错

多模态纠错





技术方案-填空题批改

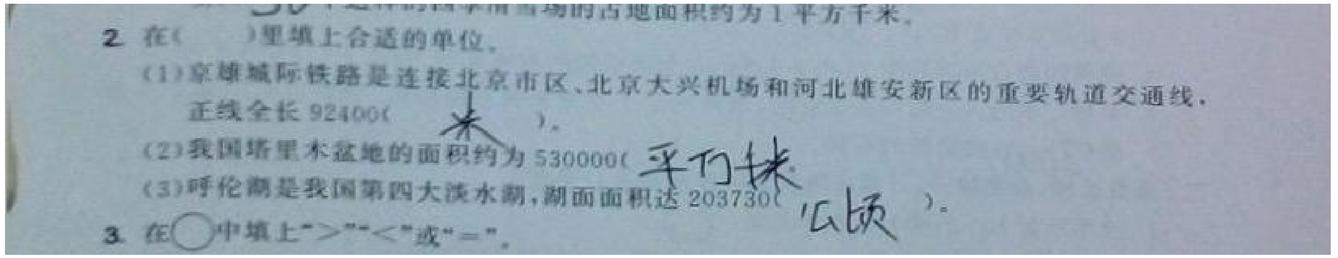
问题

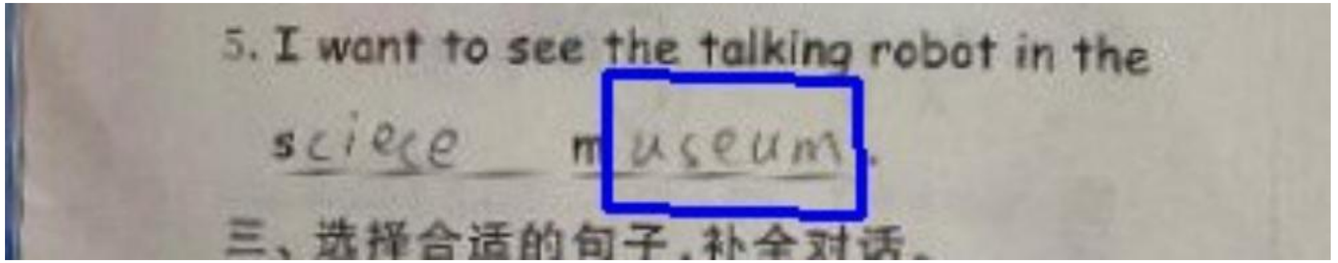
OCR错误

方案

文本纠错

多模态纠错





04

相关资料

Subject

相关资料

Improving BERT with syntax-aware local attention

Read, Listen, and See: Leveraging Multimodal Information Helps Chinese Spell Checking

Automated Essay Scoring based on Two-Stage Learning

Hierarchical Multi-task Learning for Organization Evaluation of Argumentative Student Essays.

Multimodal Machine Learning: A Survey and Taxonomy

Multi-modular domain-tailored OCR post-correction

your logo | DataFunSummit

THANKS!

今天的分享就到这里...

Ending