

2018 THINK IN CLOUD BEIJING

AI技术在教育领域的应用

李曙光 褚时科技CEO

Directory

目录

01

自动口语评测

02

自动作文评测

03

手写识别

04

AI的更多应用

01

自动口语评测

III 自动口语评测

基于GOP (Goodness of Pronunciation) :

强制对齐, 语音模型分数对比, 发现问题读音; 韵律, 语速和流利度;

加入语音识别;

深度学习: CNN, DNN, 优缺点,

移动端。

Freetalk口语评测技术:

语音分析: 对发音, 重音, 语调, 语速和流利度等方面进行分析和特征提取。

语音识别: 针对英语非母语者的语音识别, 使用深度学习方法, 识别准确率对于提取口语内容关键。

自然语言处理分析: 对识别内容在话题相关性, 语义连贯性, 语法错误, 词汇使用, 用词搭配等多维度上进行分析和特征提取。

自动打分: 多模型融合, 大量口语训练语料。

自动口语评测

应用场景：

托福和雅思口语考试自动打分和批改：目前在打分上可以取代人工，平均误差在2分左右（30分满分），大幅降低教师重复劳动。

斩托福和斩雅思：流量题库产品，覆盖80%以上的出国考生。累计百万学生使用，下面是真实产品中的口语批改报告



自动口语评测

应用场景：

学生口语日常练习：从幼儿到初高中学生的日常口语练习场景，比如单词句子跟读等，用户可以获得实时反馈，增强练习兴趣，同时降低教师检查作业的重复劳动。

大学生作业考试系统：覆盖近千所大学，下图是产品中学生口语跟读报告：

逐句朗读

| 第1句 | 第2句 | 第3句 | 第4句 | 第5句 |
|-----|-----|-----|-----|-----|
| --- | --- | --- | --- | --- |

本篇文章共分为5句，您现在朗读的是第1句。

Basic copper and cable internet access cannot run faster than about 5 megabits per second, which is just about fast enough to watch something on YouTube or for a video chat over Skype.



详情报告

口语未识别单词 口语不标准单词

Basic copper and cable internet access cannot run faster than about 5 megabits per second, which is just about fast enough to watch something on YouTube or for a video chat over Skype. But networks seldom run as fast as advertised. The cable networks upgraded with fibre can run at speeds in the hundreds of megabits per second, enough to handle most present-day applications and many in the immediate future, like high-definition video. Yet Verizon's network could leave these in the dust. In tests in Massachusetts its fibre network has run as fast as 10,000 megabits per second, and it could go faster.

02

自动作文评测

III 自动作文评测

语法错误检查：

主谓一致，动词形式的使用，词组的搭配，冠词使用、词性、选词、介词用法，动词时态等方面，Spelling，专有名词大小写和句首字母大小写等。

自然语言处理分析特征提取：

分析和统计学生作文中的文本特征，包括用词复杂度，用词搭配使用，语篇组织结构，论述连贯性和是否离题，对于议论文能否支持论点论述等等。

自动打分引擎：

针对特定的考试类型，我们使用机器学习算法（分类，排序），动态调整以上各个方面所占权重和最终的评分标准，训练相应的打分模型，最终系统可以在多个纬度给出详细的分析报告和反馈；

CNN等方法融合。

III 自动作文评测

应用场景：

初高中，托福雅思等作文自动打分和批改：取代人工打分，平均误差在2分左右（30分满分），类似ETS的e-rater；

对于小学，初中和高中学生，语法检查特别重要；

基于海量人工批改数据研发；在语法错误检测数量和精准度上远远领先于同类产品，可以和全球用户量最大的批改引擎Grammarly媲美。

当然目前我们的打分纬度除了语法，还有用词，文章脉络，句子之间连贯程度等更加复杂的方面。

自动作文评测

应用场景：

大学生作业考试系统：覆盖近千所大学， 目前已经有千万学生使用过我们的初中， 高中， 大学或者出国考试作文批改。

小学初中作业产品：覆盖千所公立中小学， 国内某大型K12培训机构。

出国考试题库系统：大幅降低教师重复劳动， 增加老师布置作文的频率。学生得到实时批改反馈， 增加练习英语写作的频率。

右图是老师使用机器批改

在托福作文教学中的批改报告：



批改报告 复制原文

名词单复数错误 3
介词错误 2
冠词错误 1
介词使用错误 1
近义词 4

This is a skeptical age, but although our faith in many of the things which our forefathers fervently believed has weakened, our faith in the benefits of knowing about events happening around the world remains the same as theirs. So, as long as the issue whether people should exert great significance in knowing about events happening around the world even some of which are unlikely that they will affect your daily life is involved, some tend to show their positive attitude towards it. However, I, with more discretion, **① crisis** that we can ill afford to determine this matter once and for all, for countries of different economic phase should be taken into account. Without doubt, by no means are we supposed to neglect the fact that in some cases, learning about events happening around the globe is helpful to develop insights and facilitate people's work, especially in highly developed countries. Considering that **② the journalism** is extremely efficient in developed nations, people have unrestrained access to all the episodes occurring on the planet. Besides, having more knowledge about other events provides people with invaluable experiences which guide people to make the most efficient choice. The example of my one of my **③ friend** in the US vividly manifests this point. My friend, Mike would stick to knowing foreign events as soon as he is free from work. He once told me that although those incidents are not likely to happen to him, he still garnered a lot of experiences **④ from** them. For example, he told me that the report on the Iraq War led him to ponder why other **⑤ nation** would like to be entangled in such a war. During this process, he drew the parallel between the Iraq War and the marketing place and helped him to save a large amount of money avoiding **⑥ competing** with other tycoons. To my certain knowledge, although those events can provide wit to people, it is irrational for us to take it for granted that it is correct under all circumstances, especially when backward countries are taken into account. On the one hand, the correspondence **⑦ in** backward **⑧ nations** is primitive and inefficient **⑨ compared** to that of the developed nations, so people have a limited resource of events around the world. On the other hand, developing basic technology is of priority in those countries and knowing about the outside world seems to distract people from achieving their original goal. The situation **⑩ of** Sudan is the case. As a backward country,

03

手写识别

III 手写识别

手写图片（扫描或者拍照）预处理和切割：文本位置识别，切行，处理纸面弯曲，行纠正，背景去噪等。

CRNN: CNN+biLSTM+CTC，行级别的End to End 建模。

MDLSTM神经网络，篇章级别End to End 建模。

语言模型在解码中很重要？

问题：书写太差，插入，断行。

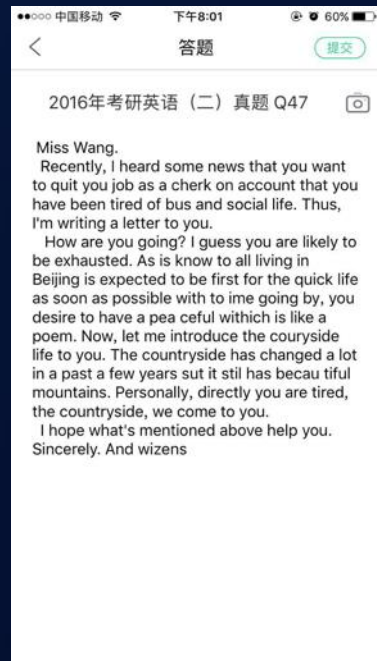
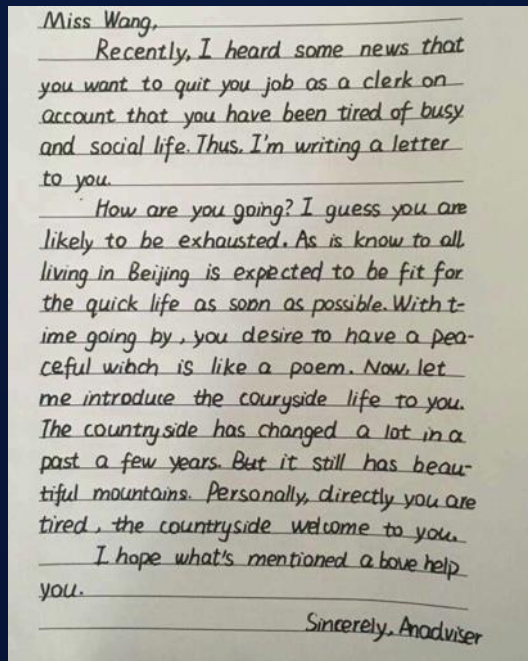
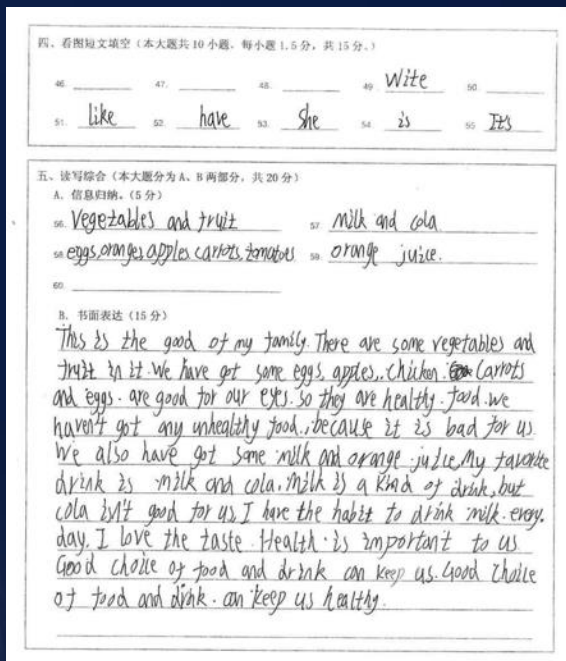
采用UCloud云服务器：训练和部署。

III 手写识别

应用场景：

产品：近千所大学考试作业系统的移动端，写作课使用。累计百万用户。雅思等纸笔考试场景的在线产品。

左一图是答题卡扫描图片；左二图是学生手写图片的照片，右一图是对应的识别结果。



04

AI更多应用

自适应学习

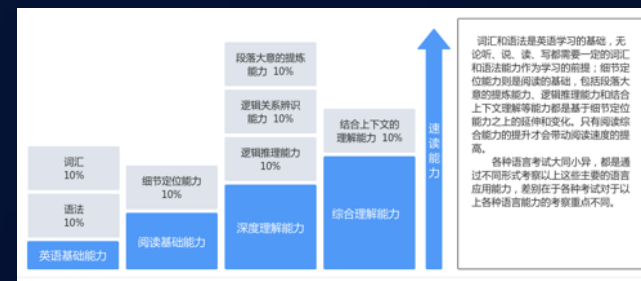
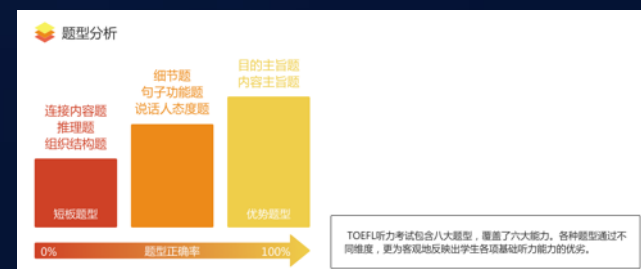
知识点建立知识图谱。

试题标签：知识点，难度，题型，考察能力等。

根据模考结果的数据分析，可以更精准的为不同能力的学生提供个性化备考计划，推送和及时调整学习路径。

自适应测试：Item Response Theory (IRT)，根据难度，区分度等建模，比较少的题目，测的更准。

右图是基于一个学生模考情况生成的能力分析结果，累计用户几十万。



III 手写板场景

以下几个领域是我们一直关注的几个方向（目前还不能很好的产品化，预计未来突破）。需要和K12培训机构或者公立学校合作。

手写公式识别：可以参考美国MyScript这个公司，已经趋于成熟。但是存在换行等问题。

数学等主观题识别：手写公式，字母识别还有汉字识别准确度提升，趋于成熟。

数学应用题解题和批改：趋于成熟。

初高中数学几何题目解题和批改：解题准确度还有提升空间，目前准确度已经超过70%；批改趋于成熟。

THANKS