

“互联网+”大学生创新创业大赛

PPT+创业计划书



DataCastle

大数据与人工智能学习平台

参赛组别：成长组

第一部分：PPT展示

第二部分：商业计划书



DataCastle

第四届“互联网+”大学生创新创业大赛

大数据与人工智能学习平台

负责人：张琳艳

指导教师：周涛 周俊临 李青

参赛成员：

张干明 周莹 张洋 赵仁杰 牛伟纳 陈瑞东 陈思宇 梅琦 庞晋雁 张季伦

DataCastle数据城堡

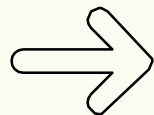
| 关于我们

公司简介

DataCastle成立于2024年，力图通过高质量课程大型算法竞赛和人才招聘服务，打造专业化的大数据与人工智能学习平台，实现人才培养和人才输出的闭环服务，一方面解决我国在相关领域高端人才稀缺和行业人才短缺的困境；另一方面提高大数据与人工智能教育在中小学的覆盖度。**是国内唯一一家通过闭环服务形式，覆盖基础教育和职业教育的学习平台。**目前注册用户超过20万，竞赛奖金累积近2000万，拥有知识产权20余项。每年收入增长超过300%，用户增长超过500%。



中小学人工智能教育
入校普及，培养种子用户



成人职业化教育人才
服务全流程管理



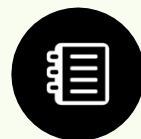
优质课程人才培养

聚集顶尖数据科学家，输出以实际应用场景为导向的优质课程，注重前沿技术和理论实践的结合。



竞赛实战人才汇聚

基于真实数据的算法竞赛，代自动编译，结果实时排名，既是高手对决的竞技场，又是学习实践的训练台。

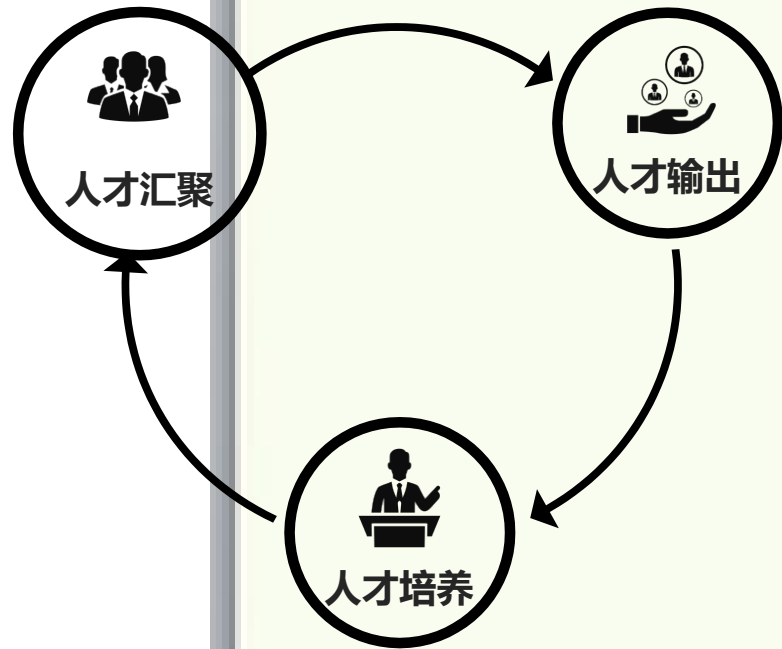


招聘服务人才输送

基于用户在线学习和竞赛记录，进行全方位画像，形成技能档案，针对岗位需求，精准推荐人才。

DataCastle数据城堡

闭环的服务体系



DataCastle

大数据与人工智能学习平台

	人才汇聚 (竞赛)	人才培养 (学院)	人才输送 (招聘)
成人	√	√	√
K12	√	√	—

数据城堡面向小学到成人的年龄阶段，打造人才汇聚人才培养人才输出的生态闭环，为用户提供大数据与人工智能竞赛学习招聘服务。

目录 CONTENT

PART 1 市场分析

PART 2 解决方案

PART 3 运营情况

PART 4 竞争优势

PART 5 财务预测





PART 1 市场分析

服务市场现状

预测：随着数据采集数据存储数据挖掘数据分析等数

我国需要更多的数据人才。

大数据人才供给高度稀缺，35年中国需要约180万数据人才，但目前

150万[1]。

《中国互联网最热职位人才报告》：**数据分析人才供**

给指数最低，仅为005，属于高度稀缺。

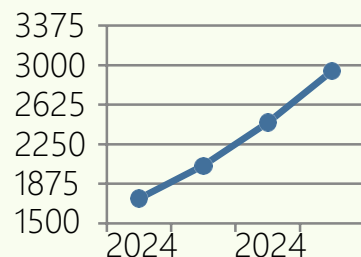
- 高校培养体系未大规模进行大数据人才培养

DataCastle业务布局目前涉及在线教育竞赛活动K12教育。

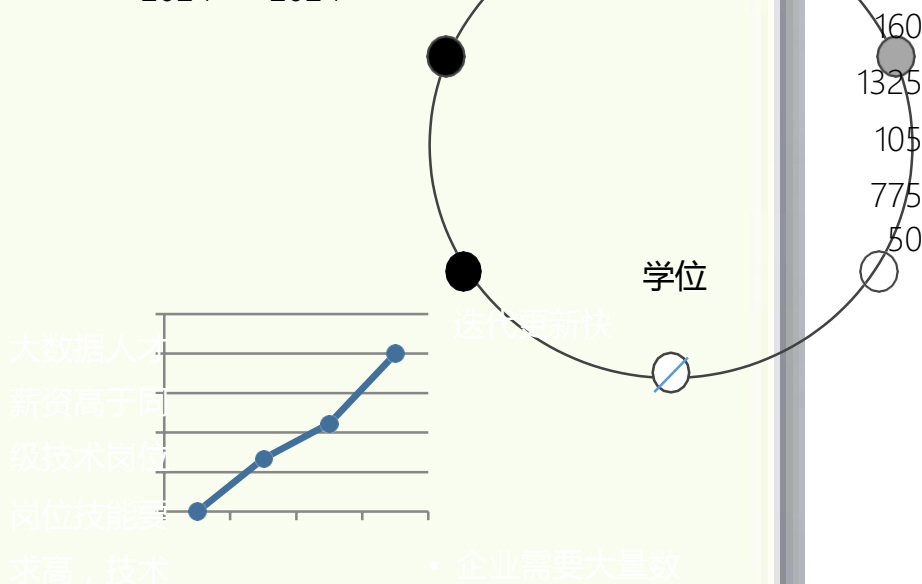
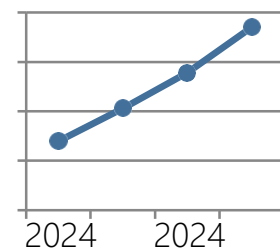
总体市场空间为千亿级，且增长速率极高。

● **成熟业务** ● **探索业务** ○ **拟开展业务** ⊗ **暂未设计业务**

单位：亿元



单位：亿元



培训需求强劲

企业端

招聘需求迫切



单位：亿元

2625
24
2175
195
1725
15

在线 教育教育

竞赛
活动

线下
培训

K12
教育

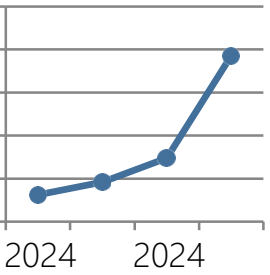
人才招聘

2375
2000
1625
1250
875

2024 2024

单
位
：
亿
元

500



数据来源：[1] 中国大数据产业峰会,清华大学计算机系教授武永卫与数联寻英联合发布的《大数据人才报告》

A photograph of three people sitting at a wooden table in a cafe or office setting. A woman in the center is wearing sunglasses and working on a laptop. A man on the left is looking at her, and a man on the right is also looking towards the laptop. The scene is dimly lit, suggesting an evening or indoor setting with low light. The text 'PART 2 解决方案' is overlaid in the center of the image.

PART 2 解决方案



在线课程

DC优选讲师，自主开发课程。涉及商务智能（BI）大数据开发数据分析数据挖掘深度学习等大数据与人工智能领域的多个分支。课程导师为知名院校的教授和专家团队业内顶级的数据科学家各行业的技术合伙人。



基于真实数据的算法竞赛，代自动编译，结果实时排名，面向全球华人；基于竞赛设计针对性的练习题目，题目支持在线批改，学员可以随时检测自己的学习效果。同时，基于课程的核心知识定制多个训练赛，学员通过分析问题提交答案，即时获得成绩和排名情况。



为普及中小学人工智能教育，K12教育遵循将教育与创客创意等相结合的理念，重在培养学生的兴趣，强调动手探索设计实验，注重知识学习过程中的思维培养，鼓励分享，培养跨学科解决问题能力团队协作能力和创新动手实践等综合素质教育。



人才招聘

基于用户在线学习和竞赛记录，进行全方位画像，形成技能档案。基于技能档案，运用大数据精准推荐算法进行人岗匹配，一方面将合适的人才推荐到合适的企业和岗位，一方面基于企业用人需求优化调整人才培养方案。

示范案例（一）

在线课程

预测型数据分析：分类及逻辑回归

二分类到多分类

One vs. Rest (OvR)

1. 为每个类别分别建立一个二分类器；
2. 训练中正例为该类别样本，负例为所有其他样本；
3. 在所有二分类器中，选择概率最高的那个类别。

使用 Logistic 进行多分类 Split Loss 默认采用 OvR 方式

```
In [27]: from sklearn import linear_model
from sklearn.model_selection import cross_val_score
import numpy as np
from sklearn.preprocessing import LabelEncoder

lm = linear_model.LinearRegression()
features = ['Petal.Length(Cm)']
X = iris[features]
y = iris['Species']
scores = cross_val_score(lm, X, y, cv=5, scoring='accuracy')
print(np.mean(scores))

~/miniconda3/envs/course_py35/lib/python3.5/site-packages/sklearn/linear_model/base.py in fit(self, X, y, sample_weight)
    480     n_jobs_ = self.n_jobs
    481     X, y = check_X_y(X, y, accept_sparse='csr', 'csr', 'csc'),
--> 482         y_numeric=True, multi_output=True)
    483
    484     if sample_weight is not None and np.atleast_1d(sample_weight).ndim > 1:

~/miniconda3/envs/course_py35/lib/python3.5/site-packages/sklearn/utils/validation.py in check_X_y(X, y, accept_sparse, y_numeric, warn_on_dtype, estimator)
    548     if y is None:
    549         assert_all_finite(y)
    550     check_consistent_length(X, y)
    551
```

LabelEncoder 我们就可以首先初始化一个LabelEncoder 的一个实例

商业模式

- 课程形式：在线课程，录播视频，用户在线购买后观看
- 运营成果：目前已超过29万人次观看学习

特色1：实训案例

DC学院基于课程的核心知识定制多个训练赛，学员需要通过分析问题提交答案，即时获得成绩和排名情况。

特色2：课后习题资料

每节课后都配套了针对性的课后练习题，课程资料区整合了每课时的核心知识，还补充了必要的基础知识以及适当的扩展内容。

特色3：学员成长曲线

系统自动记录学员每天学习时长做笔记数量答题数量答题正确率发表文章和评论的行为，以趋势图的形式展现，记录并督促学员学习。

课程观看面

示范案例（二）

| 竞赛实训

商业模式

- 合作形式：主办方提供数据奖金和运营费，DC提供竞赛运营和用户运营
- 运营成果：目前已举办超过100场竞赛，场均900支队伍，3000次结果提交

竞赛概况

[主办] 成都市人民政府

[数据] 成都市出租车GPS数据高校大学生行为数据用户多维度征信数据楼盘信息数据微博签到数据 招聘岗位数据求职简历数据创业项目数据公司注册数据项目招标数据

[奖金] 1,000,000元

[任务] 交通金融教育三大行业数据应用创新

竞赛效果

[参赛队伍] 竞赛结束时2540支，现已增至294支

[参赛人数] 竞赛结束时5350人，现已增至151人

[参赛成果] 目前吸引了来自全国科研高校和企业的一流大数据算法人员，同时也有大量海外中国留学生参与到比赛中，上线仅一个月参与人数就突破4000



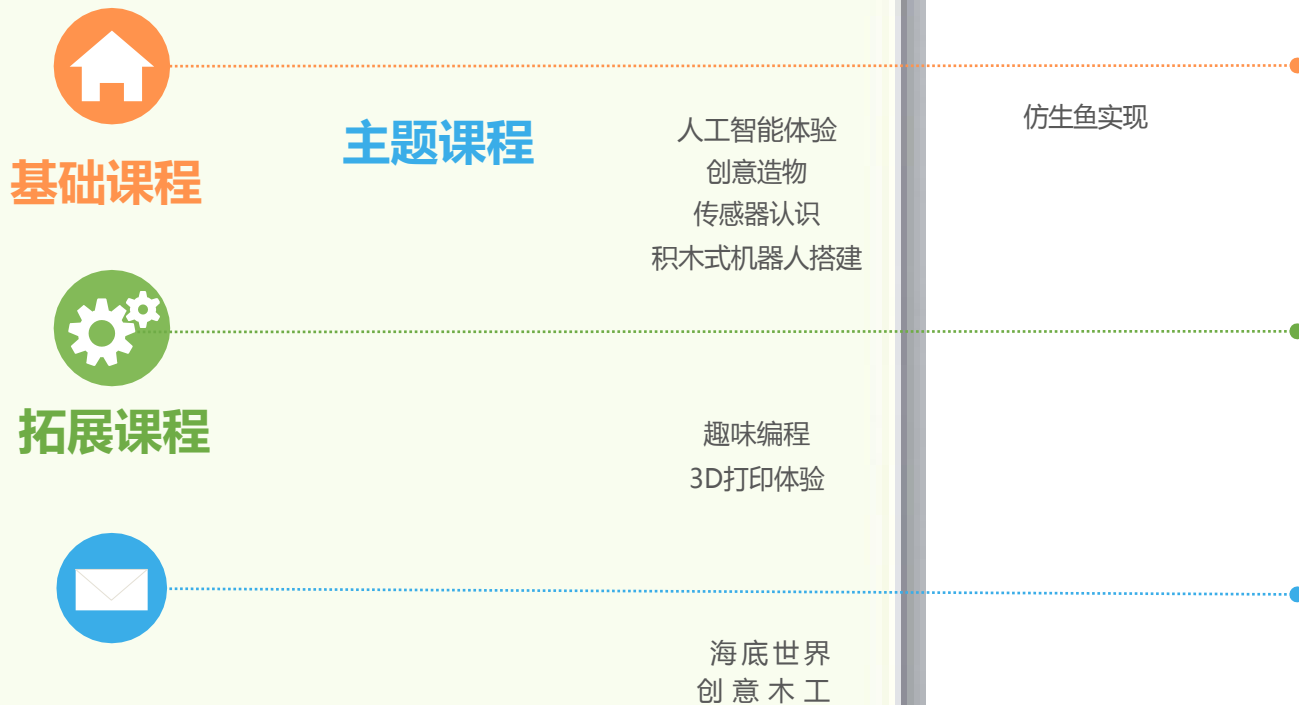
[竞赛效果] 一名选手在初赛阶段就接到数据提供企业的入职邀约两支决赛队伍先后和成都优易数据科技有限公司签订了合作协议一只决赛队伍产品受到IDG投资机构青睐

示范案例（三）

| K12教育

商业模式

- 课程形式：线下教学，有成体系的课程和配套硬件
- 盈利方式：学校/教育局集中采购课程服务





拆弹专家太阳能气象站

人工智能应用图形化编程
主题式机器人搭建

人工智能神经元程序语言编程
编程语言控制机器人
3D打印建模

人工智能探索程序语言编程
大数据时代

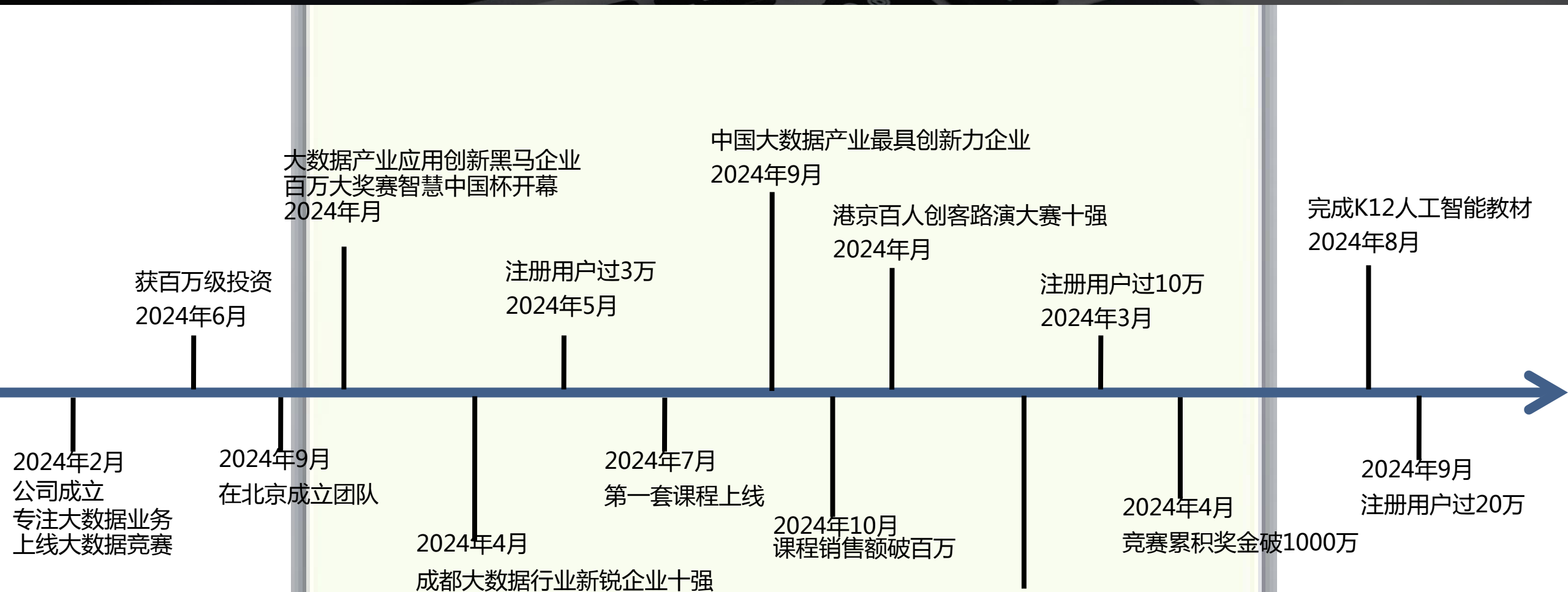
逻辑算法智能制造

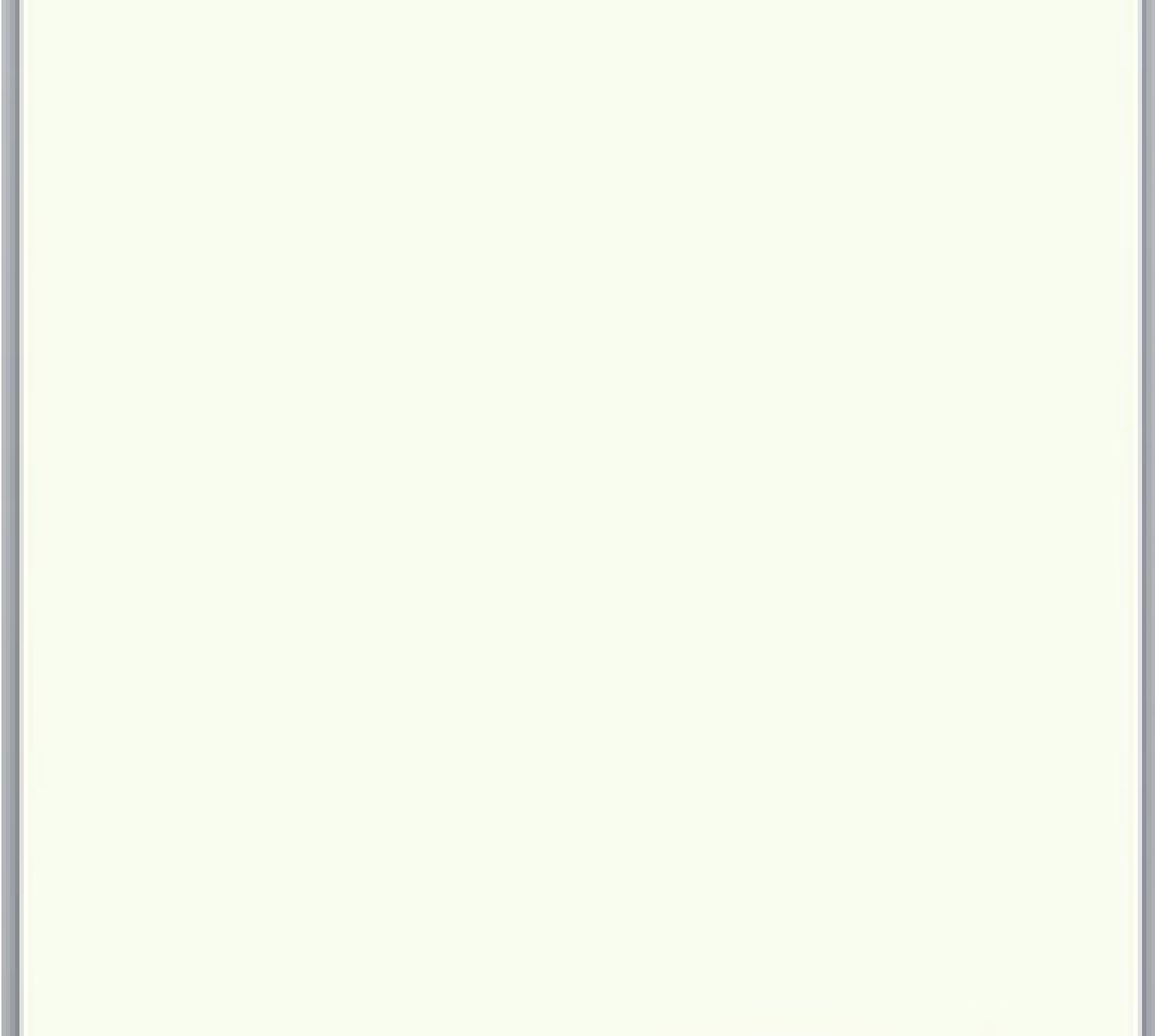
人脸识别语音识别

A top-down view of a desk with various items. In the center is a laptop with a screen showing a red circular button with the text 'START ENGINE STOP'. To the left is a tablet with a grid of app icons. In front of the laptop are two smartphones. To the right is a mouse, a pen, a small airplane model, and a cup of coffee. A desk lamp is visible in the background.

PART 3 运营情况

里程碑事件





2024年2月

入校开展K12人工智能教学

运营情况一览

在线教育

- 课程数量：200+套
- 课程观看：29万+人
- 收入总额：300万+人民币
- 单套课程最高观看人数：5万+人

竞赛实训

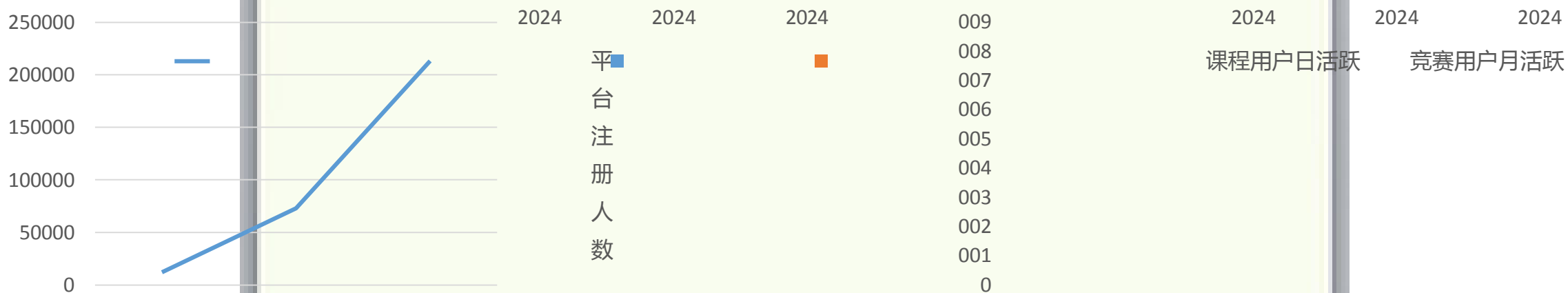
- 竞赛数量：100+场
- 比赛提交：20万+份
- 奖金总额：2000万+人民币
- 单场竞赛参赛最多人数：1万+人

K12教育

- 服务学校：50+所
- 精品课程：20+套
- 教学专家：50+人
- 师资培训：400+人

用户活跃度

- 截止2024年9月，DC平台注册用户超过20万人，年增长超过200%
- 教育平台注册用户的购课比例维持在19%上下
- 竞赛场均提交3000以上，同比增长50%



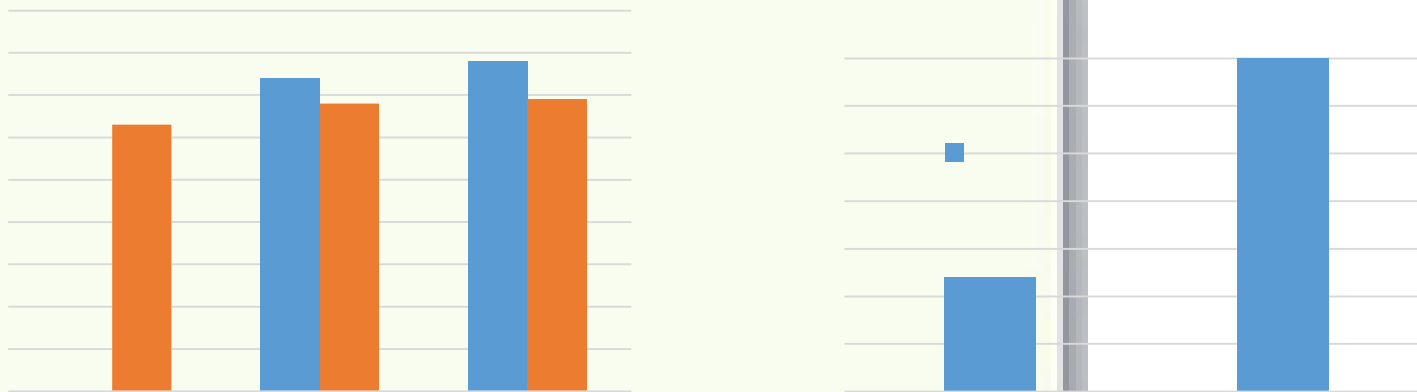


7900%
7850%
7800%
7750%
7700%
7650%
7600%
7550%
7500%

2024

2024

课程用户月活跃度



市场策略

1专业直销团队

分批在全国重点区域，按照业务种类分别搭建专业直销团队，快速占领市场份额

2线上市场运营

随着线上付费教育模式逐渐被用户接受，我们会持续增大线上用户渗透力量，通过专业自媒体渠道专业门户网站微博微信大V等进行软文推广，通过薄利多销的方式快速圈住更多用户，在原有业务上持续渗透

3. 高校渠道搭建

通过在多个高校发展校园大使，专门针对学生群体，线下推广课程和竞赛业务，形成持续用户源

4.权威品牌建设

开发对应教材参与制定行业标准汇聚海量数据，自顶向下进行专业品牌建设

A dark, artistic photograph of a workspace. In the foreground, a white cup of coffee sits on a saucer with a spoon. To the left, a laptop is open, displaying a grid of images on its screen. In the background, a person's hands are visible, writing in a notebook with a blue pen. The overall scene is dimly lit, creating a focused and professional atmosphere.

PART 4 竞争优势

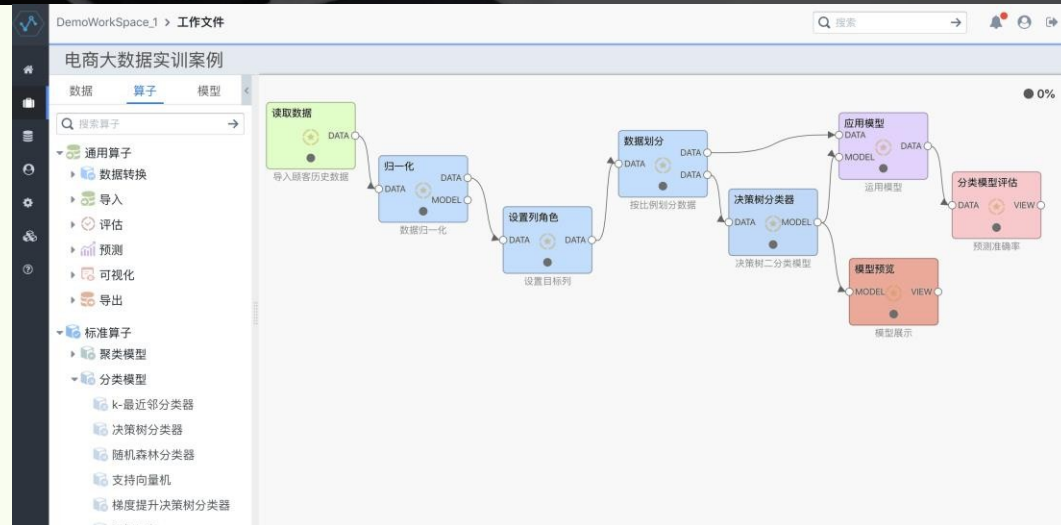
竞品分析

		课程设计	师资力量	培训模式
培 训	优达学城	覆盖数据分析机器学习无人驾驶等方向课程，纯英文，价格贵	自有讲师及10+名企工程师	线上
	DataCastle	提供系统课程，配置对应实训，让学员可进行体系化学习和成果验证	60+顶尖数据科学家知名大数据企业技术合伙人	线上+线下
		竞赛数量	客户案例	竞赛参与度
竞 赛	Kaggle	Kaggle成立于2024年，8年共举办了300+场竞赛	万事达 好事达保险 FaceBookNASA	每个竞赛平均约700支队伍参赛
	DataCastle	DataCastle成立于2024年，至今举办了超过100场比赛	CCF国家信息中心成 都市人民政府科大讯飞	每个竞赛平均约900支队伍参赛
		通用教材	课程内容	配套硬件
K12	未来创客	无通用教材	以硬件功能为主线，讲解硬件功能内容	指定硬件，成本高
	DataCastle	从小学到高中，分阶梯成体系，持续更新	以人工智能知识为主线，讲解通用知识理论和应用	通用硬件，成本低

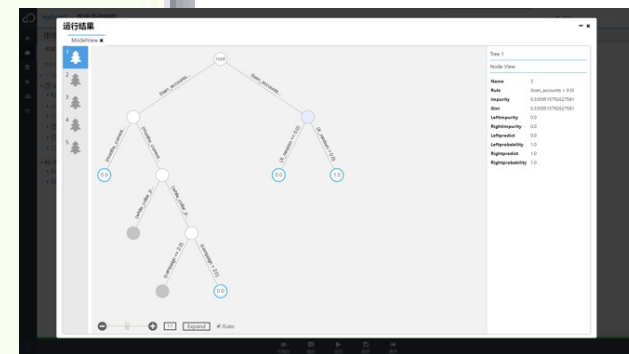
平台优势（一）

机器学习实训平台

- 1. 简单高效：**可视化的操作界面，无需编程，通过直观易懂的图形化拖拽方式即可完成数据挖掘所需的数据探索和业务建模的一系列流程，轻松驾驭大数据；随时随地可开始保存和继续进行数据价值发现的工作，大大提升数据价值发现的效率。
- 2. 多人协作：**多人协作互动地进行数据挖掘，合作伙伴之间能够随时了解数据挖掘项目的动态和进展，轻松共享建模数据和业务数据探索成果。
- 3. 功能强大：**内置了回归聚类分类问题中上百种算子，包括数据预处理数据挖掘算法模型评估算法等。
- 4. 数据安全：**支持私有云，客户不必将数据转移至第三方分析机构，内建的数据访问控制机制保证业务和数据的私密性。
- 5. 深度参与：**可开放数据价值发现的过程，提供模型查看的功能，让用户充分参



Statistics recall Result	
Options	Value
Accuracy	0.917922948073
Kappa	0.50419444471861
Reliability	0.7031024402611



与和优化业务探索的过程。

以上内容仅为本文档的试下载部分，为可阅读页数的一半内容。如要下载或阅读全文，请访问：<https://d.book118.com/135314004130011303>