



机器学习技术与平台 头豹词条报告系列

常乔雨

2023-06-28

未经平台授权，禁止转载

版权有问题？[点此投诉](#)

行业：

信息传输、软件和信息技术服务业/软件和信息技术服务业/信息系统集成和物联网技术服务

信息科技/软件服务

词条目录			
<div> 行业定义</div> <div>机器学习技术指专门研究计算机怎样模拟或实现人类的学...</div> <div>AI访谈</div>	<div> 行业分类</div> <div>机器学习可以通过学习的类型和学习算法的深度进行分类...</div> <div>AI访谈</div>	<div> 行业特征</div> <div>机器学习开发平台行业存在以下特征：1. 机器学习技术框...</div> <div>AI访谈</div>	<div> 发展历程</div> <div>机器学习技术与平台行业目前已达到 3个阶段</div> <div>AI访谈</div>
<div> 产业链分析</div> <div>上游分析 中游分析 下游分析</div> <div>AI访谈</div>	<div> 行业规模</div> <div>2022年中国机器学习与技术平台的营收规模为503.4亿...</div> <div>AI访谈 数据图表</div>	<div> 政策梳理</div> <div>机器学习技术与平台行业相关政策 5篇</div> <div>AI访谈</div>	<div> 竞争格局</div> <div>中国的机器学习技术与开发平台市场目前主要由头部互联...</div> <div>AI访谈 数据图表</div>

摘要

机器学习是一门研究计算机如何模拟或实现人类学习行为的学科，通过数学模型、算法和统计学方法分析数据，发现规律和模式，从而做出预测和决策。机器学习广泛应用于自然语言处理、计算机视觉、语音识别、机器人学、医疗保健、金融和电子商务等领域。机器学习平台是为了简化和加速机器学习模型的开发和部署过程而设计的框架或环境。目前中国机器学习框架市场呈垄断态势，其中谷歌的TensorFlow、Meta的PyTorch和百度的飞桨占据超过70%的市场份额。深度学习的高速发展是推动机器学习行业快速升级的重要因素。在2000年后，机器学习技术与平台进入了开发阶段，应用领域不断拓宽，开始广泛涉及到各类商业应用。随着数据量和计算能力的提高，机器学习平台向更加分布式、高效和自动化的方向发展。中国机器学习技术与平台行业在2015年开始进入产业化平台阶段，头部平台具有自适应性、弹性扩展性、高可用性和高性能等特点，实现了从数据清洗、特征工程到模型训练、优化和部署的全流程自动化。机器学习平台加强了与云计算、大数据、人工智能等技术的整合，进一步提升了平台的智能化水平，为企业提供了更加全面、高效的解决方案。

机器学习技术与平台行业定义

机器学习技术指专门研究计算机怎样模拟或实现人类的学习行为以获取新的知识或技能的学科。机器学习通过建立数学模型，使用算法和统计学方法分析数据，发现数据中的规律和模式，并使用这些知识来做出预测或决策。

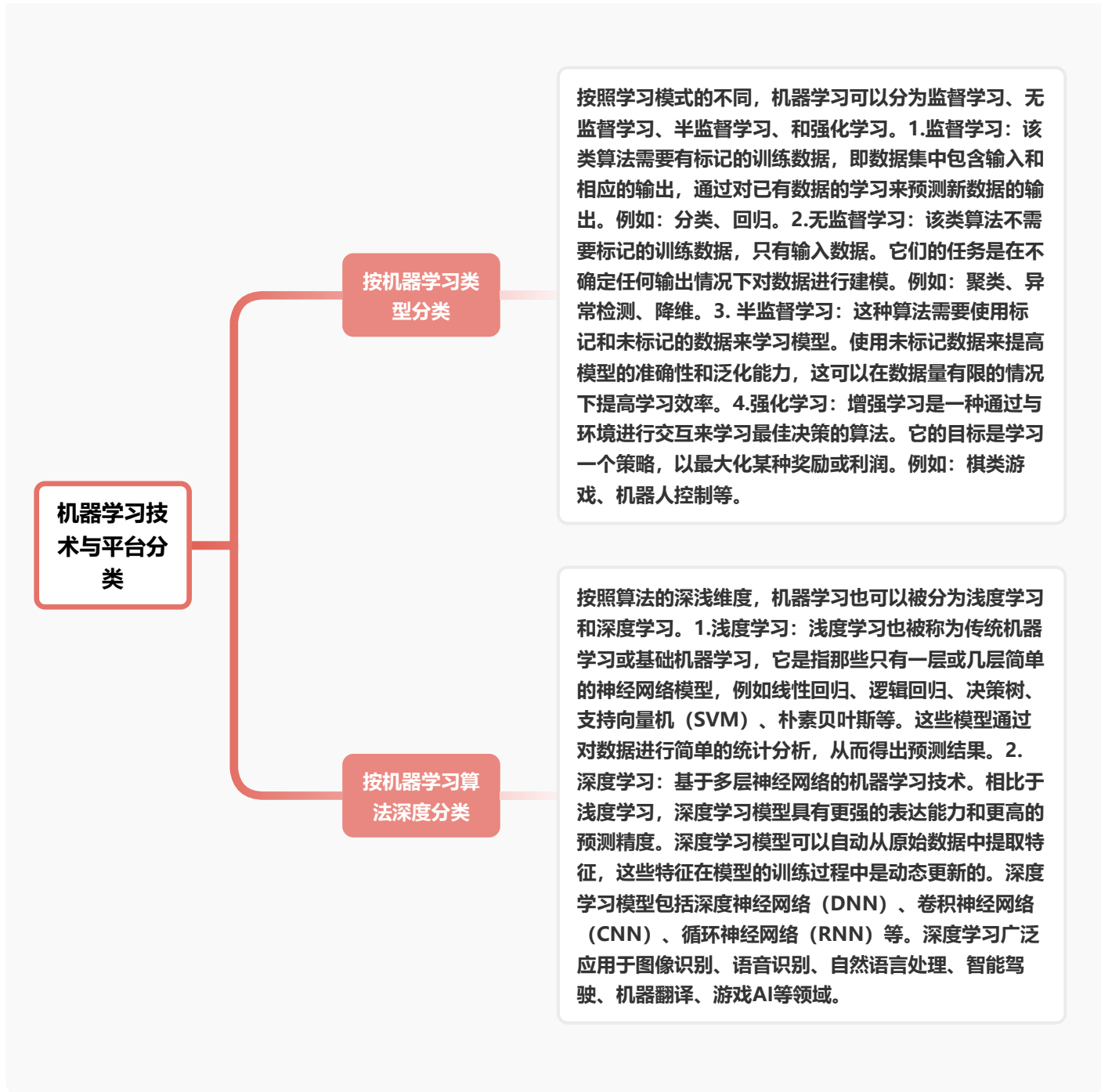
策。它可以被应用于多种领域，如自然语言处理、计算机视觉、语音识别、机器人学、医疗保健、金融、电子商务等等。机器学习平台的主要应用是为机器学习设计搭建一个框架或环境，旨在简化和加速机器学习模型的开发和部署过程。**机器学习平台提供了一种统一的环境和工具，使得数据科学家和开发人员可以在平台上进行数据预处理、特征工程、模型选择、超参数调整、和模型训练等操作，从而快速构建出高效准确的机器学习模型。**^[1]

[1] 1: 头豹研究院、清华大学...

机器学习技术与平台行业分类

机器学习可以通过学习的类型和学习算法的深度进行分类。按照学习模式的不同，机器学习可以分为监督学习、无监督学习、半监督学习、和强化学习；按照算法深浅的不同，机器学习可以分为浅度学习和深度学习。





[2] 1: 清华大学研究院、头豹...

机器学习技术与平台行业特征

机器学习开发平台行业存在以下特征：1. 机器学习技术框架呈垄断态势；2.深度学习助推机器学习快速发展；3. 支持向量机技术为机器学习领域最热门研究领域。

1 机器学习技术框架呈垄断态势

中国机器学习框架的市场呈垄断态势。中国目前市场前三大机器学习框架为谷歌的TensorFlow，Meta的PyTorch，和百度的飞桨稳居前三，三者共占据超70%的市场份额。

TensorFlow、PyTorch和飞桨是中国机器学习市场占据超70%的前三大框架。它们得到广大开发者的青睐，原因在于其卓越的性能、强大的社区支持和高度易用性。三者均使用行业内顶级的计算性能，能够在大规模数据集上进行快速训练和调整。此外，它们都有强大的开源社区支持，提供了大量的技术支持、学习资源和代码实例等。易用的API和丰富的文档使得用户能够快速入门和进行深度学习实验。以上三大框架在市场竞争中占据垄断地位，为广大AI开发者提供了强有力的工具支持。

2 深度学习助推机器学习快速发展

深度学习凭借其适应性强、准确率高等特点获得市场高度关注。应用领域不断拓宽。深度学习的高速发展是推动机器学习行业快速升级的重要因素。

深度学习的优势主要体现在图像、语音、以及人脸的识别方面。图像识别方面，深度学习可通过叠加海量数据和计算机并行运算能力显著提升图像识别精确度，将识别的错误率降低至2.3%。在语音识别方面，深度学习可以显著降低语音识别的误差率，错误率可降低至5%以内，实用价值明显。在人脸识别方面，通过学习大量的人脸数据，现阶段的人脸识别技术准确率已达99.9%以上，广泛应用在各类身份识别场景。综上所述，深度学习通过在图像识别、语音识别、人脸识别等技术领域的技术突破广泛应用在各大领域，助推机器学习技术与平台高速发展。

3 支持向量机技术为机器学习领域最热门研究领域

通过AMiner对超2亿篇论文的数据进行深入挖掘，支持向量机（SVM）为当前热度最高的机器学习研究领域

领域技术分析系统利用AMiner超过2亿篇论文的数据进行深入挖掘，对技术趋势、国际趋势、机构趋势以及学者趋势等方面进行分析。根据技术趋势分析，我们可以发现，该领域当前最热门的话题是支持向量机（SVM）。从全局热度来看，机器向量机在2007年的热度以指数形上升，在2018年开始登上热度榜首。

[3] 1: <http://trend.amine...> 2: 清华大学研究院、头豹...

机器学习技术与平台发展历程

在2000年以前，机器学习平台主要集中在研究实验室和大学内部，由于计算资源和算法库有限，应用受限于研究和实验目的。在2000年后，随着技术的成熟度提高，机器学习技术与平台进入了开发阶段。行业从理论研究向应用转变，开始广泛涉及图像处理、语音识别、自然语言处理、推荐系统等领域，并开始涉及到各类商业应用。随着数据量的增大和计算能力的提高，机器学习平台开始向更加分布式、高效和自动化的方向发展。许多大型科技公司也开始提供云计算平台来处理大规模数据，标志着行业进入了平台开发与建立的发展阶段。**2015**

年开始，中国机器学习技术与平台行业进入产业化平台阶段，通过打造成成熟的商业解决方案为不同行业提供定制化的机器学习服务和解决方案。此发展阶段的头部平台具有自适应性、弹性扩展性、高可用性和高性能等特点，实现了从数据清洗、特征工程到模型训练、优化和部署的全流程自动化。在这一阶段，机器学习平台加强了与云计算、大数据、人工智能等技术的整合，进一步提升了平台的智能化水平，为企业提供了更加全面、高效的解决方案。^[4]

[5]

- 2000年以前属于机器学习技术与平台行业的发展初期阶段。**机器学习平台主要集中在一些研究实验室和大学内部，开发者使用基于Python和MATLAB等编程语言的算法来实现机器学习模型。**但由于计算资源和算法库有限，这些平台往往只能应用于研究和实验目的。**

在2000-2010年期间，机器学习平台的发展特征主要是从理论研究向应用转变。这个阶段的机器学习应用开始涉及到更广泛的领域，如图像处理、语音识别、自然语言处理、推荐系统等，同时也开始涉及到更多的商业应用。此外，随着数据量的增大和计算能力的提高，机器学习平台开始向更加分布式、高效和自动化的方向发展。例如Google发布的分布式计算框架Hadoop和MapReduce和微软发布的Azure云计算都为机器学习处理海量数据提供了优质的解决方案。**许多大型科技公司也开始提供云计算平台来处理大规模数据，并提供机器学习库和工具，如Google的TensorFlow、Microsoft的Azure ML和Amazon的SageMaker等，标志着行业进入平台开发与建立的发展阶段。**

5/20

- 在产业化平台阶段，机器学习平台进一步发展为成熟的商业解决方案，面向不同行业提供了定制化的机器学习服务和解决方案，如金融、医疗、交通等。这一阶段的机器学习平台具有自适应性、弹性扩展性、高可用性和高性能等特点，可以实现从数据清洗、特征工程到模型训练、优化和部署的全流程自动化。例如百度的飞桨，阿里云的PAI，和华为的Model Arts都是行业内优秀的机器学习平台，为企业提供了商业化的解决方案。此外，机器学习平台在这一阶段还加强了与云计算、大数据、人工智能等技术的整合，进一步提升了平台的智能化水平，为企业提供了更加全面、高效的解决方案。

4: 清华大学研究院、南京...

6/20

器学习的整体应用情况在下游呈现井喷势态。然而，上游的供给能力在当下呈现出供给不足的态势。机器学习上游供给短缺主要表现在数据和人才两个方面。数据方面，企业和组织存在数据收集和管理困难，数据质量不足或不完整，同时对于一些敏感领域的的数据，如医疗和金融数据，隐私和法规问题也限制了数据的共享和使用。人才方面，机器学习需要高度技术化的人才进行算法研究、模型设计、数据处理和系统部署等工作，但机器学习领域的人才非常稀缺，特别是具备深度学习和神经网络等前沿技术的人才更为缺乏。据报告显示，全球现在已经面临100万人才缺口，到2020年将攀升至226万。虽然全国有367所高校具备人工智能研究方向，但每年毕业的学生仅有2万人，远远不能满足市场的需求。这种供给短缺问题导致许多企业和组织在实践中遇到困难，例如在医疗领域的图像识别方面，由于数据质量不足和人才缺乏等问题，无法建立准确和高效的图像识别系统。此外，机器学习人才的供给差异也是问题之一，一些欠发达国家或地区的机器学习人才非常短缺，加剧了全球机器学习上游供给短缺的问题。因此，机器学习产业需要加强对数据和人才的投入，以缓解上游供给不足的问题，促进产业链的健康发展。^[6]

上 产业链上游

[7]

生产制造端

机器学习基础层

上游厂商

腾讯云计算（北京）有限责任公司 >

百度云计算技术（北京）有限公司 >

阿里云计算有限公司 >

查看全部 v

产业链上游说明

产业链上游的机器学习基础层主要包含为机器学习技术与平台提供的基础设施层与研究机器学习创新的高精尖人才。基础设施层包括数据采集与处理、芯片设计、算法开发、以及必要的硬件设备制造。机器学习工程师使用硬件设备和芯片，对获取的大量高质量数据进行分析 and 运算从而产出优质的算法以生产出对应的行业解决方案。以下内容将对机器学习平台行业上游的各细分环节进行深入探讨。从数据采集与处理环节分析，机器学习中的数据治理生命周期包括确定数据来源和采集方式、考虑数据质量和获取访问权限、数据清洗去除无效数据并解决数据冲突、数据集成整合不同来源和格式的数据以便后续分析、数据转换包括将数据转换成统一格式、进行特征提取和构建新特征、数据存储和管理需将处理好的数据存储到数据库或文件系统中并进行管理和维护，以便于后续访问和使用。这些步骤对机器学习模型的性能和准确性具有重要影响，需要按照生命周期的流程进行数据采集和处理以确保数据的质量和准确性，提高机器学习模型的性能和应用效果。中国面向AI的数据治理市场规模增速迅猛，景气度高，根据相关数据统计，中国面向AI的数据治理市场目前为30亿元，未来五年内有望以28%的年增速继续高速增长。从芯片设计环节分析，目前适用于机器学习的主流芯片有图形处理器（GPU）、张量处理器（TPU）、可编程门阵列（FPGA）、应用集成电路（ASIC）、以及神经处理单元

(NPU)。不同类型的机器学习芯片适用于不同的应用场景，在选择特定芯片时需要考虑场景的特点需求，例如性能、功耗、定制化和可扩展性。**机器学习芯片在中国的本土产业化程度较低，2022年我国GPU的行业本土化不足10%，FPGA的行业本土化不足20%。**中国在芯片制造的技术理论和关键设备仍落后于国际一流水准。未来，随着中央对机器学习硬件本土化的高度重视以及投资，中国会逐步拉近与国际芯片制造水平的差距。**从算法设计环节分析，机器学习算法可按技术操作角度的划分为监督、非监督、以及强化学习；也可以按具体算法分为三大类：聚类算法、分类算法、以及回归算法。**机器学习核心算法包括：1.回归算法：线性回归、最近邻。回归算法是对数值型连续随机变量进行预测和建模的监督学习算法。案例包括机场客流量分布预测、房价预测、股票走势等数值连续变化的场景。2.分类算法：逻辑回归、决策树、支持向量机、朴素贝叶斯。分类算法是对离散型随机变量建模或预测的监督学习方法。案例包括邮件过滤、预测学生通过率、贷款违约率等预测类别的场景。3.聚类算法：K均值聚类、主成分分析、层次聚类。聚类算法是基于数据的内部结构寻找关注样本的自然集群的无监督式学习算法。案例包括基于用户信息的行业选址、电商产品自动推送、检测疾病等围绕相似点聚类集群的场景。**机器学习算法开发是机器学习平台上游的重要因素组成，算法的优劣性和工程师对算法的选择和适配会直接影响机器学习平台对下游解决方案产出的质量。**算法工程师也是目前行业内薪资最高的职位之一，根据企业调研，阿里、华为、百度等一线算法工程师的年薪高达45万-200万元之间，算法工程师的应届生平均起薪为32,280元，由此可见算法对于机器学习平台的重要性以及企业对人才重视程度。**未来，算法作为机器学习平台的核心上游产出之一，不断提升算法工程师的技能和能力，对于推动机器学习平台的发展和企业的竞争力至关重要。**

中

产业链中游

品牌端

机器学习技术与平台

中游厂商

- 阿里云计算有限公司 >
- 深圳市亚马逊云计算科技有限公司 >
- 百度云计算技术（北京）有限公司 >
- 查看全部 ▾

产业链中游说明

机器学习技术与平台的中游由机器学习的技术框架和行业解决方案供应平台组成。机器学习技术框架是一种软件架构，旨在简化机器学习模型的开发和部署。它提供了一组工具和库，使得机器学习开发者可以更快地构建、测试和部署模型。主流的技术学习框架包含TensorFlow、Pytorch、Scikit-learn、百度飞浆等。框架在不同的场景有着差异化的优势，开发者可根据自身的需求选择适合的框架。行业解决方案平台则为企业和机构提供了一站式机器学习解决方案的平台。平台通过提供各种工具、技术和部署方案，帮助企业通过机器学习赋能。**从机器学习技术框架的市场竞争力角度分析，中**

国机器学习框架的市场呈垄断态势。中国目前市场前三大机器学习框架为谷歌的TensorFlow，Meta的PyTorch，和百度的飞桨稳居前三，三者共占据超过70%的市场份额。以上三大机器学习框架之所以成为众多开发者的首选的主要原因有三点：**1. 性能卓越：**三大平台均使用了行业内顶级的计算性能，使得它们能够在大规模的数据集上进行快速地训练和整合。Tensorflow在CPU的运行速度上为所有机器学习框架的第一名，而pytorch则在GPU的运算速度方面傲视群雄。**2. 社区支持：**三大平台均有强大的开源社区支持。据统计，2019年，69% 的 CVPR 论文、75% 以上的 NAACL 和 ACL 论文，以及 50% 以上的 ICLR 和 ICML 论文都选择使用 PyTorch。知名社区Github上有关Pytorch和Tensorflow的帖子数量超过65%。强大的社区支持为用户提供了大量的技术支持、学习资源、代码实例等，使得用户的使用黏性逐步提高。**3.易用性高：**三大平台都提供了易用的API和丰富的文档，使得用户可以快速入门并进行深度学习实验。此外，它们还提供了丰富的预训练模型和工具，进一步降低了使用门槛。极高的易用性让以上平台成为众多AI软件开发者的首选。**综上所述，中国机器学习框架在市场竞争中的垄断地位得益于其卓越的性能、强大的社区支持和高度易用性，为广大AI开发者提供了强有力的工具支持。从一站式机器学习解决方案平台的市场竞争力角度分析，中国机器学习平台的第一梯队由各大互联网巨头带领，分别为亚马逊、百度、商汤、华为、腾讯、以及阿里云。第二梯队包括深耕机器学习平台开发多年的厂商，包括第四范式、火山引擎、旷世科技等拥有丰富人工智能领域研发经验和垂直行业部署的潜力增长者。在第一梯队中，百度依靠着行业领先的研发投入和早期的布局目前在机器学习的完整性、开放性、和前瞻性均处在领先地位。根据百度2021年第四季度的财报披露，AI业务收入达到了98.4亿元。AI业务的营收占比达全公司收入的34%，为行业之最。根据相关数据统计，百度的机器学习框架百度飞桨以及机器学习平台EasyDL在中国的市场份额均为第一名。**

下

产业链下游

渠道端及终端客户

机器学习技术与平台的应用场景

渠道端

中国国际金融股份有限公司 >

中国中信集团有限公司 >

华泰证券股份有限公司 >

查看全部 ▾

产业链下游说明

在大数据时代，机器学习技术与平台已成为众多行业数字化转型的必然之选。随着机器学习技术的不断发展，它的使用已广泛应用于金融、医疗、零售、互联网、政务等各类行业。**从下游应用的发展趋势角度分析，机器学习将深入渗透到各行各业，成为继互联网革命之后下一个改变产业模式的革命性技术产品，其划时代的跨行业特质主要体现在以下几个要素：1.自动化：**机器学习技术能够帮助各行

业自动化许多任务，替代昂贵的人力，达到降本增效的目的。例如，对于制造业，机器学习可以通过自动化生产线上的工作，减少人为的错误和成本，提高生产效率和质量。对于金融业，机器学习可以自动化风险评估、投资决策和交易执行等任务，提高效率和准确性。Varma的一项研究表明当一家企业使用一台智能语音机器人时，可帮助客服团队省去330个小时/月的工作量，几乎等于两个员工一个月的工作时间。

2.预测性分析：机器学习可以分析大量的数据，从中提取出规律和趋势，并做出未来的预测，大大提高预测的准确性，为各行各业带来赋能。例如，对于金融银行业，机器学习可以利用历史数据预测欺诈的发生和传播情况，帮助银行降低因欺诈行为导致的财产损失。蚂蚁金服曾发表论文表示基于Attention及LSTM的GeniePath学习网络关键路径，在实际防骗保应用的准确率可以达到95%。

3.个性化服务：机器学习可以分析个人的偏好和行为，从而提供个性化的服务。例如，在旅游业中，机器学习可以分析消费者的旅行历史和偏好，为其提供个性化的旅行服务。在教育业中，机器学习可以分析学生的学习经历和兴趣，为其提供个性化的学习体验。友盟+与达摩院联合开发了国内首个智能推送功能，它可以针对不同的推送对象和文案进行精准匹配，通过感知不同用户场景和各种配额限制，选择最佳的分配策略。该功能已经得到了广泛的应用，并在某电商APP客户中取得了显著的效果，客户推送消息的点击率提升了35.87%。

从机器学习技术与平台在下游场景的细分应用分析，机器学习在互联网、金融业、以及零售业的渗透率最高，均超过60%。通过深入探究机器学习在以上行业内的应用，可以发现以下几个要点：

1.在互联网行业，推荐系统是互联网行业中应用机器学习最广泛的领域。推荐系统通过分析用户的历史行为数据，如购买记录、搜索记录等，来预测用户的偏好并向其推荐个性化的商品、服务、信息等。例如，亚马逊的推荐系统会根据用户购买历史和浏览记录，向用户推荐类似的商品。

2.在金融行业，风险控制和投资决策是最重要的两个应用领域。在风险评估方面，金融机构可以利用机器学习模型来评估借款人的信用风险和违约风险，识别准确率可超过95%，可以更好地管理贷款组合。在资产管理方面，机器学习技术可以帮助投资者分析市场数据和经济趋势，以做出更好的资产配置决策。著名私募对冲基金桥水和Citadel通过使用机器学习的方法去做策略配置，在2022年全球经济整体停滞的一年均获得超100亿美元的利润，成为了2022年全球私募对冲基金营收前两名，Citadel也成为了历史上最盈利的对冲私募基金。

3.在零售行业，机器学习两大核心的应用场景是定价和预测库存需求。定价方面，机器学习技术可以分析市场供需、竞争对手价格等因素，预测最佳的定价策略，并帮助零售商优化价格策略，以实现最大利润。例如，某科技咨询公司曾利用机器学习模型为一家零售企业选取了一系列变量因素，包括促销折扣幅度、促销时间、无促销情况下的平均销售额、等来进行未来的销售预测。相比与该零售企业的传统预测方法，基于机器学习的预测方法将该零售企业的预测准误差率从35%降低至24%，大幅提高了预测不同场景的消费准确率，帮助企业更好地制定定价策略。

预测库存需求方面，好的库存管理可以最大限度地利用资金，提高企业效率，做出最经济合理的决策。例如，某电商公司库存中的大部分商品销量不佳，只有10%的商品被销售，部分商品一年中只销售一次。因此缩减库存和预测销量对该商家的重要性不言而喻。通过分析历史销售记录，包括销售记录、库存数量、市场营销类型、产品上架次数、产品上架时间和价格等因素，结合ABC分类控制管理，使用Sample算法将数据分为训练集和测试集，再综合多种机器学习算法，得到新的库存管理系统。经过测试，该系统的预测精度高达80%，明显优于传统库

存管理的40%，证明该系统可以更好地预测商品需求和最佳库存，从而帮助该电商公司提高效率和降低成本。

[6] 1: <https://www.fxbao...> 2: 头豹研究院、华为、安...

[7] 1: <https://www.jiqizhi...> 2: <https://www.jiqizhi...> 3: <https://www.jiqizhi...> 4: <https://finance.sin...> 5: <https://blog.csdn....> 6: <https://www.sohu....> 7: <https://www.jiqizhi...> 8: <https://leaddesk.c...> 9: <https://info.umen...> 10: <http://www.360d...> 11: <http://yongsheng...> 12: 清华大学研究院、阿里...

机器学习技术与平台行业规模

2022年中国机器学习与技术平台的营收规模为503.4亿元，过去五年市场规模的复合增长率达到52.5%。预计，未来数年内中国机器学习与技术平台的市场规模将继续稳步增长，在2026年达到1,158.3亿。

在过去的数年中，机器学习技术与平台的市场规模在中国呈现快速增长的趋势。其中有两点原因对机器学习技术与平台的市场规模产生了强力助推作用。

1. 数字化转型加速：随着中国经济的快速发展和科技的不断突破，数字化转型在近年来成为各类行业企业竞争的主旋律。越来越多的企业开始关注机器学习的技术应用，以达到提高生成效率、降低运营成本的目的。机器学习技术是企业数字化转型的核心，它可以帮助企业从大量数据中提取有价值的信息和洞察，进而优化业务流程、降本增效、和为产品服务带来创新。根据相关数据统计，机器学习应用的虚拟客服技术每处理一次对话可节省5.5美元的成本。应用机器学习技术为企业带来的收入增长直接贡献率平均可以达到6.3%。因此，企业对机器学习的重视程度越来越高。

2. 政策大力支持：近年来，中国政府一直致力于支持机器学习的发展。例如，2017年中国政府发布了《新一代人工智能发展规划》，旨在通过投资、政策支持和技术创新等手段，推动中国成为全球领先的人工智能创新中心。国家发展和改革委员会在2018年宣布将向人工智能产业投入1,000亿元人民币的专项资金。2019年，国务院印发了《新一代人工智能发展规划（2018-2020年）》，其中提出了多项政策措施，包括加强人才培养、建设公共服务平台等。

综上所述，数字化转型加速和政策大力支持是中国市场规模快速增长的主要原因。

预计在未来五年内，中国机器学习与技术平台将会继续保持高速增长态势，市场规模在2026年达到1,158.3亿，CAGR达到18.1%。在政策的助推下，机器学习技术在中国不断拓展应用场景，市场规模将持续高速增长。

1. 机器学习的应用场景在不断拓展：机器学习在中国的应用场景正在不断扩大。从传统的金融、医疗、交通等领域到新兴的智能家居、智能制造、智能物流等领域，机器学习技术正逐渐成为这些领域中的核心技术之一。根据相关数据统计，机器学习在政务的渗透率有望在未来五年内从47%提升至90%以上，工业领域的应用将从30%提升至50%以上。

2. 政策的正面影响还在继续：国务院明确指出人工智能核心产业到2030年将会达到一万亿元。2022年，人工智能核心产业的市场规模达到5,000亿，预计未来7年内还有较大的增长空间。

因为机器学习技术

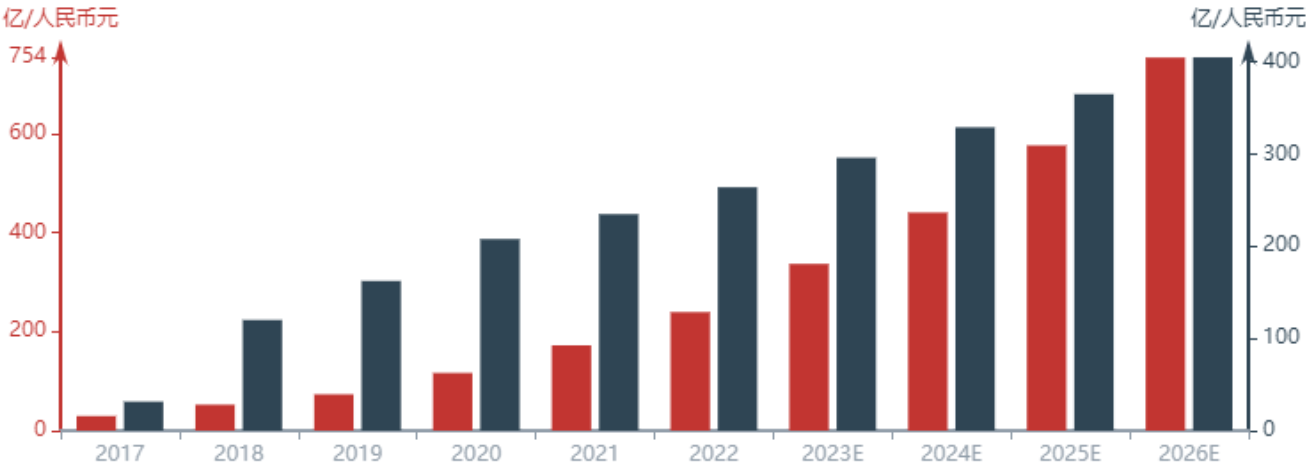
应用场景不断拓展，且政策利好的影响仍在继续，预计未来五年内中国机器学习与技术平台市场规模将保持高速增长，呈现出非常乐观的前景。^[8]

中国机器学习平台行业行业规模

第四范式招股书、弗若斯特沙利文、头豹研究院

下载原始数据

中国机器学习技术与平台行业市场规模



机器学习技术与平台市场规模=机器学习技术市场规模+机器学习平台市场规模

[8] 1: <https://www.voach...> 2: <https://www.ligon...> 3: <https://www.voach...> 4: 国家发改委、IBM、第...

机器学习技术与平台政策梳理

[9]	政策	颁布主体	生效日期	影响
	《国家创新驱动发展战略纲要》	国务院	2016-05	8
政策内容	发展新一代信息网络技术，增强经济社会发展的信息化基础。加强类人智能、自然交互与虚拟现实、微电子与光电子等技术研究，推动宽带移动互联网、云计算、物联网、大数据、高性能计算、移动智能终端等技术研发和综合应用，加大集成电路、工业控制等自主软硬件产品和网络安全技术攻关和推广力度。			
政策解读	纲要明确提出要加强人工智能基础研究，探索人工智能技术应用的前沿领域，重点支持人工智能领域的科技创新。建设人工智能高水平创新团队，推进人工智能人才国际化合作，为人工智能的发展提供有力的人才保障。推进人工智能与实体经济深度融合，加快人工智能在各领域的应用，为产业升级和经济转型提供支持。总体来说，纲要对人工智能的支持主要围绕技术研发、人才培养、应用推广和法律制度等方面展开，旨在推动人工智能技术的创新和应用，促进经济社会的发展。			

政策性质	指导性政策
------	-------

[9]	政策	颁布主体	生效日期	影响
	《新一代人工智能发展规划》	国务院	2017-07	9
政策内容	提出了面向2030年我国新一代人工智能发展的指导思想、战略目标、重点任务和保障措施，部署构筑我国人工智能发展的先发优势，加快建设创新型国家和世界科技强国。到2020年，人工智能总体技术和应用与世界先进水平同步，人工智能产业成为新的重要经济增长点，人工智能技术应用成为改善民生的新途径。			
政策解读	《新一代人工智能发展规划》是中国政府在人工智能领域的一项重要规划，强调要推动新一代人工智能在各领域的应用，包括工业、交通、医疗、教育、金融、安防等方面，以实现经济社会的高质量发展。明确提出构建开放协同的人工智能科技创新体系。旨在大力推动中国人工智能应用的规范和管理，保障人工智能发展的安全和可持续性。			
政策性质	规范类政策			

[9]	政策	颁布主体	生效日期	影响
	《数据安全管理办法》	国家互联网信息办公室	2019-05	6
政策内容	对于数据出境安全的管理做了更为详细的规定，明确了境内数据出境的审查程序和出境安全评估标准。通过国家强制力对隐私信息的收集使用予以限制，在隐私信息泄漏时亦有迹可循，以实现个人隐私信息的数据安全。			
政策解读	机器学习在进行训练时需要大量的数据，而这些数据可能涉及到个人隐私和商业机密等敏感信息，因此数据的安全处理和出境安全评估非常重要。该办法明确了数据出境的管理程序和安全评估标准，有助于加强对机器学习领域的出境安全管理。数据是机器学习的基础，加强对数据的保护对机器学习平台产业有着重要的利好影响。			
政策性质	指导性政策			

[9]	政策	颁布主体	生效日期	影响
	《国家新一代人工智能标准体系建设指南》	国家标准化管理委员会、中央网信办	2020-07	7

政策内容	提出到2021年，明确人工智能标准化顶层设计，研究标准体系建设和标准研制的总体规则，明确标准之间的关系，指导人工智能标准化工作的有序开展，完成关键通用技术、关键领域技术、伦理等20项以上重点标准的预研工作。到2023年，初步建立人工智能标准体系，提供公共服务能力。
政策解读	指南规划了新一代人工智能标准体系的总体框架和具体内容，包括标准目录、标准体系结构、标准分类和标准制定程序等。通过建设完备、系统、规范的人工智能标准体系，促进人工智能技术的创新和应用，保障人工智能的安全和可持续发展。建设内容提出规范AI研发与应用和数据存储，包括数据治理、共享开放等核心要素。
政策性质	指导性政策

[9]	政策	颁布主体	生效日期	影响
	《工业物联网创新发展行动计划 (2021年-2023年) 》	国家工信部	2021-02	8
政策内容	工业互联网是新一代信息通信技术与工业经济深度融合的全新工业生态、关键基础设施和新型应用模式。它以网络为基础、平台为中枢、数据为要素、安全为保障，通过对人、机、物全面连接，变革传统制造模式、生产组织方式和产业形态，构建起全要素、全产业链、全价值链全面连接的新型工业生产制造和服务体系。			
政策解读	该政策明确表示计划通过三年的时间推进工业互联网基础设施、产业应用和创新发展，加快工业数字化转型，构建数字经济新生态。提出了一系列具体的措施，如推进工业互联网标准化、构建工业互联网开放平台、培育重点产业领域的工业互联网应用示范等。重点任务指出支持工业5G芯片模组、边缘计算专用芯片与操作系统等人工智能基础元件。			
政策性质	鼓励性政策			

[9]	1: http://www.gov.cn/zhe...	2: 国家工信部、国务院、...	3: http://www.gov.cn/zhe...	4: https://baijiahao.baidu...
	5: https://mp.weixin.qq.c...	6: http://www.gov.cn/xin...		

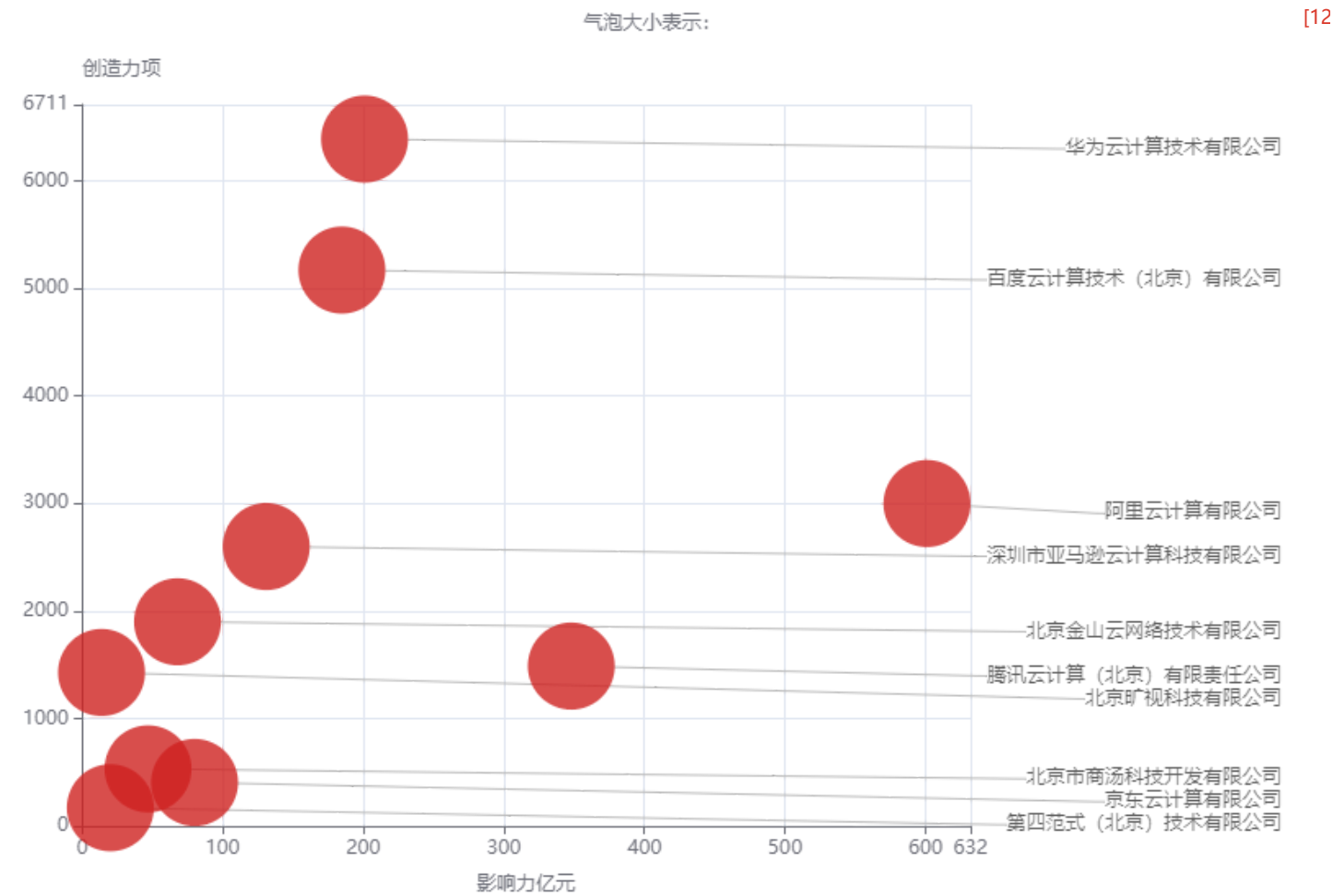
机器学习技术与平台竞争格局

中国的机器学习技术与开发平台市场目前主要由头部互联网科技大厂组成。主要的竞争者有：1. 头部互联网科技公司：腾讯云、百度智能云、华为云、商汤科技、阿里云、亚马逊云、京东云。2.垂直领域创新机器学习技

术与开发平台：旷世科技、第四范式、创新奇智。目前市场份额排名前列的企业有百度云、阿里云、腾讯云、和 华为云。

中国机器学习技术与开发平台行业的参与者少，行业准入门槛高，竞争者多为头部互联网科技大厂。这是由于机器学习开发具有极高的技术门槛和昂贵的成本，需要具备高水平的技术实力才能保证平台的稳定性、性能和安全性。例如，像OPENAI这样的领头企业，需要耗费近20年的时间研究和打磨ChatGPT，并且拥有来自世界顶尖科研机构的科研团队。机器学习技术与开发平台的投资成本也是极高的，百度在机器学习领域的总投入超过1,000亿元，核心研发投入占核心收入比例连续多个季度超过20%。

中国机器学习平台竞争格局在未来数年将会维持稳定，主要参与者仍将由百度云、阿里云、腾讯云、华为云、亚马逊云、微软云、商汤科技等互联网科技大厂组成。虽然新兴垂直领域中小型厂商可以通过垂直竞争能力来保持一定的市场份额，但整体来看无法撼动头部厂商的领先地位。在头部厂商中，百度云以233亿的投入金额排名第一，第二名则是阿里云，2022年的投入金额为151.5亿。因此，在机器学习技术与开发平台稳定的头部竞争格局中，百度云较有希望从众多大厂中脱颖而出。然而，随着机器学习技术的不断发展和应用场景的不断扩大，未来的竞争格局仍可能发生变化。新兴技术公司和创新型企业可能会挑战现有的头部互联网科技大厂，尤其是在一些新兴领域和应用场景中。**因此，对于机器学习技术与开发平台企业来说，除了加强技术研发和提高技术实力外，还需要不断关注市场变化和技术趋势，积极拓展新的应用场景和业务领域，以保持竞争优势和领先地位。**^[10]



横坐标代表企业的“影响力”：以竞争主体企业平台的综合营收能力体现，通过规模大小一定程度体现企业的市场影响力。位置越靠右侧，企业的营收能力越强。纵坐标代表“创造力”：以竞争主体在人工智能领域的专利拥有数量体现，位置越靠上方，机器学习平台企业的潜在创造力越强。

上市公司速览

[10]

1: <http://stock.hexun...>

2: <https://baijiahao.b...>

3: <https://www.megv...>

4: <https://docs.ksyun...>

5: <https://www.huaw...>

6: <https://cn.aliyun.c...>

7: <https://cloud.tenc...>

8: <https://xueqiu.co...>

9: 中国科学院、企业官网...

[11]

1: <https://cloud.ofwe...>

2: <https://www.huaw...>

3: <https://cloud.tenc...>

4: <http://stock.hexun...>

5: <https://baijiahao.b...>

6: 企业官网、企业招股书...

[12]

1: <https://baijiahao.b...>

2: <https://xueqiu.co...>

3: 中国科学院、企查查、...

机器学习技术与平台代表企业分析

1

竹间智能科技（上海）有限公司

[13]

^

公司信息

企业状态	存续	注册资本	9000万美元
企业总部	市辖区	行业	科技推广和应用服务业
法人	简仁贤	统一社会信用代码	91310115350689047D
企业类型	有限责任公司(港澳台法人独资)	成立时间	2015-08-25
品牌名称	竹间智能科技（上海）有限公司		
经营范围	从事智能科技、网络科技、计算机科技、信息科技领域内的技术开发、技术咨询、技术转让... 查看更多		

竞争优势

1. 应用领域广泛、为多个行业的头部企业进行服务：竹间智能的产品广泛应用于电商、金融、物联网、教育、零售、旅游、能源等垂直行业中。竹间智能曾为交通银行、招商银行、民生银行、太平洋保险、华夏人寿、恒大地产等多个行业龙头企业进行产品服务。

竞争优势2

2. 技术深耕垂直行业：竹间智能深耕深度学习、自然语言处理、情感识别等技术。并且，根据客户特定的需求提供基于人工智能技术的定制化服务，以定制化服务深入应用场景帮助企业打造具备垂直行业知识的AI机器人。

2 北京第四范式智能技术股份有限公司

公司信息

企业状态	存续	注册资本	45124.3288万人民币
企业总部	市辖区	行业	软件和信息技术服务业
法人	戴文渊	统一社会信用代码	91440300311653228T
企业类型	股份有限公司（外商投资、未上市）	成立时间	2014-09-17
品牌名称	北京第四范式智能技术股份有限公司		
经营范围	技术开发、技术服务、技术咨询；应用软件服务；数据处理（数据处理中的银行卡中心，P... 查看更多		

竞争优势

1.技术研究能力优秀：首席执行官为华为诺亚方舟实验室主任科学家，首席科学家是国际人工智能协会院士并担任国际人工智能联合会理事会主席，首席科学家曾任今日头条推荐系统负责人。

竞争优势2

2. 产品应用场景丰富：第四范式的产品可以应用于银行、保险、政务、能源、智能制造、零售、医疗、证券等领域，落地应用案例超过一万个。

3 云从科技集团股份有限公司【688327】 [14]

公司信息

企业状态	开业	注册资本	74067.0562万人民币
企业总部	广州市	行业	软件和信息技术服务业
法人	周曦	统一社会信用代码	914401153314442716
企业类型	其他股份有限公司(上市)	成立时间	2015-03-27
股票类型	科创板	品牌名称	云从科技集团股份有限公司
经营范围	计算机系统服务;人工智能行业应用系统集成服务;人工智能通用应用系统;人工智能基础资源... 查看更多		

财务数据分析

[14]

财务指标	2017	2018	2019	2020	2021
资产负债率(%)	23.377	66.95	34.269	39.777	56.141
营业总收入同比增长(%)	-	650.171	66.768	-6.512	42.494
应收账款周转天数(天)	-	100	114	185	140
流动比率	5.546	1.59	2.839	2.542	1.869
每股经营现金流(元)	-0.79	-2.98	-0.84	-0.74	-0.87
毛利率(%)	36.7893	21.7024	40.8874	-	-
流动负债/总负债(%)	74.108	89.726	87.653	79.684	83.854
速动比率	5.282	1.452	2.724	2.314	1.669
摊薄总资产收益率(%)	-	-18.429	-33.763	-33.485	-26.318
加权净资产收益率(%)	-96.5	-38.39	-342.55	-	-
基本每股收益(元)	-1.43	-2.04	-2.91	-1.31	-1.01
净利率(%)	-191.8905	-41.3385	-85.7251	-111.8133	-61.7492
总资产周转率(次)	-	0.494	0.426	0.311	0.448
每股公积金(元)	-	-	3.0133	4.1053	4.3878
存货周转天数(天)	-	68	75	110	106
营业总收入(元)	6453.37万	4.84亿	8.07亿	7.55亿	10.76亿
每股未分配利润(元)	-	-	-1.2692	-2.519	-3.5271
稀释每股收益(元)	-1.43	-2.04	-2.91	-1.31	-1.01
归属净利润(元)	-106316310.66	-180675218.51	-639601351.64	-812933299.48	-632128411.39
扣非每股收益(元)	-1.51	-2.38	-3.19	-1.4	-1.2

经营现金流/营业收入	-0.79	-2.98	-0.84	-0.74	-0.87
------------	-------	-------	-------	-------	-------

▪ 竞争优势

1. 人脸识别技术优势强：云从科技是一家专注于人脸识别领域的公司，凭借多年的技术积累，已经取得了在该领域的重大突破。除了人脸识别技术外，该公司还在图像识别、语音识别、自然语言处理、机器学习等领域拥有领先的技术实力。云从科技自主研发了多项人工智能技术，例如跨镜追踪（ReID）、3D 结构光人脸识别、双层异构深度神经网络和对抗性神经网络技术，这些技术均处于业界领先水平。

▪ 竞争优势2

2. 服务生态能力强：专注于构建 AI 应用与服务生态体系的公司，致力于打造 AI 平台型厂商。该公司的人机协同操作系统已经具备全链的 AI 技术整合能力和高效的 AI 生产力。在“AI 三浪论”的背景下，市场供需关系向着有利于深耕垂直领域、打造平台化 AI 厂商的发展方向转变。因此，云从科技作为平台化 AI 厂商，有着良好的发展前景。

▪ 竞争优势3

3. 金融领域渗透率高：在金融领域，云从科技已经成功服务了超过100家银行总行平台，占有市场接近83%的份额。该公司的核心竞争优势在于抓住了头部客户，并且这些头部客户对全链技术的应用需求不断提高，单个大客户订单规模也在逐渐增加。随着供需关系的变化，大客户和大项目为云从科技提供了机会，使得该公司能够始终站在行业信息化趋势最前沿，形成了深厚的认知壁垒。在这个基础上，云从科技进一步优化数据模型，实现了“感知、认知、决策”闭环，在二浪时代中形成了领先的地位。

[13] 1: 云从科技招股书、云从...

[14] 1: 竹简智能科技官网、头...

法律声明

权利归属：头豹上关于页面内容的补充说明、描述，以及其中包含的头豹标识、版面设计、排版方式、文本、图片、图形等，相关知识产权归头豹所有，均受著作权法、商标法及其它法律保护。

尊重原创：头豹上发布的内容（包括但不限于页面中呈现的数据、文字、图表、图像等），著作权均归发布者所有。头豹有权但无义务对用户发布的内容进行审核，有权根据相关证据结合法律法规对侵权信息进行处理。头豹不对发布者发布内容的知识产权权属进行保证，并且尊重权利人的知识产权及其他合法权益。如果权利人认为头豹平台上发布者发布的内容侵犯自身的知识产权及其他合法权益，可依法向头豹（联系邮箱：support@leadleo.com）发出书面说明，并提供具有证明效力的证据材料。头豹在书面审核相关材料后，有权根据《中华人民共和国侵权责任法》等法律法规删除相关内容，并依法保留相关数据。

内容使用：未经发布方及头豹事先书面许可，任何人不得以任何方式直接或间接地复制、再造、传播、出版、引用、改编、汇编上述内容，或用于任何商业目的。任何第三方如需转载、引用或基于任何商业目的使用本页面上的任何内容（包

括但不限于数据、文字、图表、图像等），可根据页面相关的指引进行授权操作；或联系头豹取得相应授权，联系邮箱：support@leadleo.com。

合作维权：头豹已获得发布方的授权，如果任何第三方侵犯了发布方相关的权利，发布方或将授权头豹或其指定的代理人代表头豹自身或发布方对该第三方提出警告、投诉、发起诉讼、进行上诉，或谈判和解，或在认为必要的情况下参与共同维权。

完整性：以上声明和本页内容以及本平台所有内容（包括但不限于文字、图片、图表、视频、数据）构成不可分割的部分，在未详细阅读并认可本声明所有条款的前提下，请勿对本页面以及头豹所有内容做任何形式的浏览、点击、引用或下载。