



浪潮信息

SOLUTION

DNA基因组模式识别人工智能
解决方案



01

应用背景

PROJECT BACKGROUND

DNA研究主要通过DNA的对比研究，来比较两类不同个体或群体之间遗传的差异和联系。通过识别分析出基因组数据一些特定的DNA或基因序列片段或者一些突变序列的演进等，辅助探索在疾病研究、大型绝灭哺乳动物的迁移演化、动植物家养驯化过程及生物与环境的相互关系等诸多方面的未知奥秘，并对人类的社会生活模式和环境作用等产生一定的启示。浪潮信息为DNA研究提供了异构加速人工智能计算节点，助力其通过深度学习的方法从海量基因数据中识别出新基因，从分子学机制角度加速DNA研究，不断为开拓新的领域。

02

业务挑战

BUSINESS CHALLENGE

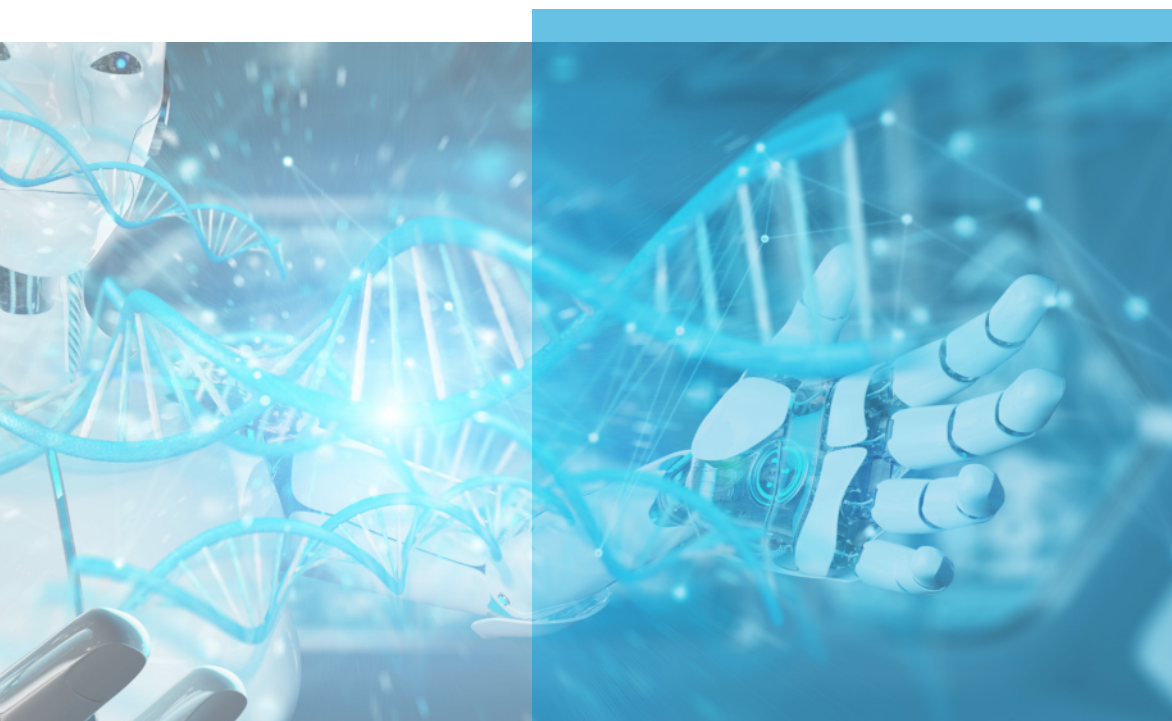
- DNA容易降解或损伤，在DNA保存的过程中也容易受到寄居微生物的“入侵”，测序过程中也常被“污染”，除加强测序控制手段外，测序分析需要尽可能高效地识别并剔除污染序列。
- 数据和规模的难题：基因测序会用高通量的测序方法产生大量的测序数据。基因组数据库数据非常庞大，特定或未知基因的识别并非易事，同时对计算设备的计算、内外存储性能都形成了较高挑战。
- 人工智能深度学习模型复杂，AI算法计算效率优化、软硬件调优及业务流程验证跑通在生命科学专家看来很棘手。

解决方案

SOLUTION

03

- 将基因测序数据和基因组数据数值文本化，用最先进的卷积神经网络构建深度学习算法模型，由机器推荐给出特定序列或未知序列，并结合人工科学分析，去除污染序列，找到识别未知或特定序列，兼顾人工智能的高效和科学研究的准确。
- 采用MLPerf冠军性能产品浪潮信息NF5488A5人工智能服务器，提供640GB的GPU全局共享显存，NVSwitch八卡全速互联，可承载最大规模的全基因组模式识别深度学习应用，并具备强大异构算力快速完成模型训练和推理识别。
- 部署人工智能深度学习管理系统AIStation，内置容器化仓库镜像，秒级部署模型开启训练，同时AIStation推理版可快速推理识别，一键跑通AI业务。



方案亮点

SOLUTION HIGHLIGHTS

04

- 架构领先：提供最先进的NVSwitch全互联人工智能服务器NF5488A5，解决基因组大数据训练难题。
- 平台成熟：采用 AIStation 人工智能平台软件，用户方便易用，可快速的开展 AI 模型训练的相关研究,并为应用加速提供了良好的服务支撑,更好的开展研究业务。
- 提质增效-AI辅助科研：创新科研模式，通过AI高效识别基因序列，大幅降低工作繁复性，并通过科学分析验证，整体有效降低研究耗时，提升效率50%以上

现代人祖先和尼安德特人的交流基因智能识别、癌基因智能识别等



DNA基因识别model

DNA基因组模式识别人工智能平台



NF5488A5

基因训练数据



客户收益

CUSTOMER BENEFITS

05



- 使用浪潮信息NF5488A5人工智能服务器配合浪潮信息整体人工智能解决方案，DNA研究可从我国边疆地区的DNA基因库中更快、更准地识别出我国北方地区的草原游牧民族，如匈奴、东胡、鲜卑、乌桓、契丹、蒙古等诸族的特定基因，并基于此对人群间的迁徙及混杂的过程进行研究，从全新的角度论证我国的民族融合史和社会文明起源变迁，增强民族自信，助力华夏复兴。

浪潮信息电子信息产业股份有限公司 www.ieisystem.com

浪潮信息技术支持与服务热线 400-860-0011

购买咨询热线 400-860-6708 / 800-860-6708

文中有关产品图片及文字仅供参考，详细产品规格及价格，请向浪潮信息当地经销商查询。版权声明©2023浪潮信息，保留一切权利。BJ 06/2023



浪潮信息服务号