

AI4HEP数据集 20220811

1 “具有对称保持注意网络的无置换多射流事件重建”数据集

2021 Permutationless Many-Jet Event Reconstruction with Symmetry Preserving Attention Networks

作者: Michael James Fenton 一作+通讯

期刊: Physical Review D IF5.41 SCI二区 引用1

单位: 加州大学欧文分校

用对称保持注意网络进行**无置换多射流事件重建**。

大型强子对撞机产生的顶夸克具有复杂的探测器特征, 需要特殊的重建技术。

最常见的衰变模式是全射流”all-jet”通道下, 会产生6-jet的终态, 由于在pp碰撞中存在大量排列组合, 特别难以重建。

使用广义注意力机制, 发展了对称保持注意网络(Symmetry Preserving Attention Networks, SPANet), 网络可以在不进行组合爆炸(combinatorial explosion)的情况下明确识别出每个顶夸克的衰变产物, 明显优于现有SOTA方法, 分配所有的射流all-jets, 对6-jet 7-jet 8-jet事件的准确率分别为80.7% 66.8% 52.3%。

网络: Transformer编码+tensor attention

源码: SPANet <https://github.com/Alexanders101/SPANet> torch框架

数据集: 2021_ttbar数据集, http://mlphysics.ics.uci.edu/data/2021_ttbar/ 大约5GB

2 通过深度学习有效检测反物质物理中的反氢

2017 Efficient Antihydrogen Detection in Antimatter Physics by Deep Learning

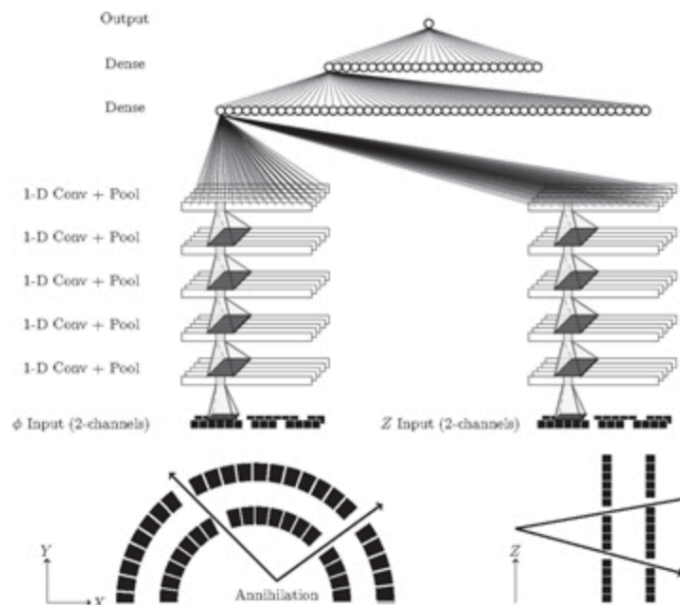
作者: P Sadowski

期刊: Journal of Physics Communications 新期刊 CS2.6 引用7次

单位: 加利福尼亚大学计算机系

蒙特卡洛模拟和退火时间重建

网络: 5层1D卷积, 2个FC层



数据集: antihydrogen数据集 <http://mlphysics.ics.uci.edu/data/antihydrogen/> 包含200万个事件

无代码

3 深度学习增强的Higgs玻色子到 τ - τ 搜索

2015 Enhanced Higgs Boson to Tau Tau Search with Deep Learning

作者: P. Baldi

期刊: PHYSICAL REVIEW LETTERS, PRL, IF9.19 CS16.8 SCI一区, 引用64次

单位: 加州大学欧文分校物理与天文学系

希格斯玻色子提供质量, 目前的分析技术缺乏跨越5sigma限制性障碍的统计能力。神经网络用来检测希格斯玻色子到tau-tau轻子对的衰变, 发现的意义相当于累计数据集25%增量。

网络: 8层非线性结构深度网络 SPEARMINT 贝叶斯优化算法

数据集: htautau <http://mlphysics.ics.uci.edu/data/htautau/>

无代码

问: 5sigma significance barrier是什么?

4 高能物理中射流子结构的深度神经网络分类

2016 Jet Substructure Classification in High-Energy Physics with Deep Neural Networks

作者: Pierre Baldi

期刊: PRD IF5.41 SCI二区 引用130次

单位: 加利福尼亚大学计算机系

大型强子对撞机中, 产生大量高速粒子, 强子衰变过程是准直的从而产生射流重叠jets overlap。观察jet的子结构是由单个粒子引起的还是由多个衰败物体引起的是重要的问题。传统方法: 专家模式检测热量计中的能量沉积。收集二维图像。该网络可以与当前最先进的区分单个强子粒子喷流和成对准直强子粒子重叠配流的方法想比较, 或略优。

网络结构: 神经元100-500, 全连接层1-5, 局部链接0-5

图像分类任务

数据集: hepjets <http://mlphysics.ics.uci.edu/data/hepjets/>

5 高能物理的参数化神经网络

2016 Parameterized neural networks for high-energy physics

P. Baldi

期刊: EUROPEAN PHYSICAL JOURNAL C, IF4.99 CS8.3 SCI二区 引用85次

研究了神经网络分类器, 输入测量特征, 物理参数, 应用与高能物理问题。

应用: 理论模型参数 (如粒子质量) 进行参数化的工具, 该工具允许单个网络在一系列质量上提供改进的区分。

数据集: HEPMASS dataset <http://mlphysics.ics.uci.edu/data/hepmass/>

6 通过深度学习寻找高能物理中的奇异粒子

2014 Searching for Exotic Particles in High-Energy Physics with Deep Learning

作者: P. Baldi

期刊: Nature Communications IF17.69 CS23.3 SCI一区 508次引用

分类度量能力提高8%, 可以提高对撞机寻找奇异粒子的能力

代码: <https://github.com/uci-igb/higgs-susy> 56颗星

数据集:

HIGGS dataset <http://mlphysics.ics.uci.edu/data/higgs/>

SUSY supersymmetry dataset: <http://mlphysics.ics.uci.edu/data/susy/>

7 基于深度神经网络的高能物理喷注味分类

2016 Jet Flavor Classification in High-Energy Physics with Deep Neural Networks

期刊: PRD 引用88次

数据集: Jet Flavor dataset http://mlphysics.ics.uci.edu/data/hb_jet_flavor_2016/

8 利用高斯过程推断中微子振荡参数的有效方法

2020 Efficient Neutrino Oscillation Parameter Inference Using Gaussian Processes

期刊: PRD 引用1次

数据集 neutrino <http://mlphysics.ics.uci.edu/data/neutrino/>

9 学习鉴别电子

2021 Learning to Identify Electrons

Baldi通讯

PRD 引用两次

研究了在对撞机实验中通常用于区分电子与喷流背景的最先进分类特征是否忽略了有价值的信息。对电磁和强子量热计沉积物的深度卷积神经网络分析与典型特征的性能进行了比较,揭示了约5%的差距,这表明这些较低级别的数据确实包含未开发的分类能力。为了揭示这些未使用信息的本质,我们使用了一种最新开发的技术,将深度网络映射到物理可解释的可观测空间。我们确定了两个简单的量热计观测值,它们通常不用于电子识别,但模拟了卷积网络的决策,几乎缩小了性能差距。

数据集: 2020_electron: http://mlphysics.ics.uci.edu/data/2020_electron/

10 基于解耦生成模型的稀疏图像生成

2019 Sparse Image Generation with Decoupled Generative Models

Baldi

数据集: muon_2020 http://mlphysics.ics.uci.edu/data/muon_2020/

11 学习分离 μ 介子

2021 Learning to isolate muons

期刊: JOURNAL OF HIGH ENERGY PHYSICS IF6.38 CS10.3 SCI二区

在对撞机物理数据分析中,区分重玻色子衰变产生的快速 μ 子和与重味喷流产生相关的 μ 子是一项重要任务。我们探讨了量热计沉积物中是否存在标准隔离锥方法无法捕捉的信息。我们发现,访问量热计单元的卷积网络和粒子流网络的性能超过了隔离锥的性能,这表明 μ 子周围量热计沉积物的径向能量分布和角度结构包含未使用的分辨能力。我们收集了一小部分高级观测值,这些观测值汇总了量热计信息,并通过直接分析量热计单元的网络缩小了性能差距。这些观测值在理论上定义良好,可以用对撞机数据进行研究。

网络: 全连接和卷积浅层

数据集: 2021_muon http://mlphysics.ics.uci.edu/data/2021_muon/

12 SARM:粒子物理中可伸缩生成稀疏图像的稀疏自回归模型

2021 SARM: Sparse Autoregressive Model for Scalable Generation of Sparse Images in Particle Physics

Baldi通讯

网络: 自回归网络

模拟数据的生成对于粒子物理中的数据分析至关重要,但当前的蒙特卡罗方法计算成本非常高。基于深度学习的生成模型已经成功地以较低的成本生成了模拟数据,但在数据非常稀疏的情况下很难生成。我们引入了一种新的深度稀疏自回归模型(SARM),该模型以

可处理的可能性显式学习数据的稀疏性，与生成对抗网络（GAN）和其他方法相比，使其更加稳定和可解释。在两个案例研究中，我们将SARM与GAN模型和非解析自回归模型进行了比较。作为性能的定量度量，我们计算在生成图像和训练图像上计算的物理量分布之间的瓦瑟斯坦距离（W-p）。在第一项研究中，以90%像素为零值的射流图像为特征，SARM生成的图像的W-p分数比其他最先进的生成模型获得的分数高24%-52%。在第二项研究中，在 μ 子附近的量热计图像上，其中98%的像素为零值，SARM生成的图像的W-p分数高出66%-68%。使用其他度量进行的类似观察证实了SARM对于粒子物理中稀疏数据的有效性。

数据集 2020_SARM http://mlphysics.ics.uci.edu/data/2020_SARM/

冒出来一本书：Artificial Intelligence for Science Alok Choudhary 2023Jane 新书
DOI: <https://doi.org/10.1142/13123>