

# “华为杯”第五届中国研究生人工智能创新大赛赛题

单位:中国研究生人工智能创新大赛组委会

发布时间: 2023-04-28

来源: 中国研究生人工智能创新大赛

阅读次数: 7017

## 一、报名及提交作品要求

参赛队伍在以下赛题中任选其一进行作品创作参赛。7月31日前，在大赛官网报名，通过培养单位资格审查后，8月7日前，在官网提交参赛作品。（作品提交规范及模板见附件）

大赛官网 <https://cpipc.acge.org.cn/> (<https://cpipc.acge.org.cn/>)

## 二、大赛赛题

（一）技术创新：开放题

（二）应用创意：开放题

（三）企业赛题：华为赛题共3个，均为华为技术有限公司从实际需求出发拟定的与智能技术密切相关的题目。属于此种选题方式的作品还可以参与华为专项奖的评奖。

### 1. 华为赛题一

#### 1.1 题目名称

AI交互预测决策规划

#### 1.2 题目描述

城区场景，拓扑复杂且存在车辆、非机动车、行人等复杂交互目标，期望通过AI和大数据学习的方法预测环境并给出类人安全的决策规划解，算法层面满足环境预测和决策规划的自治性和时序稳定性。基于白盒算法的决策规划方法，存在价值函数定义困难、调参难度大、类人性差的问题；基于AI的预测决策规划方法存在模型黑盒可解释性差、结果可靠性无法保障的问题。如何基于数据驱动的方法实现安全可靠的预测决策规划算法存在以下挑战：

1. AI模型建模拓扑和交通参与者，需要实现对world model的有效建模；
2. AI模型设计需要充分考虑自车和社会交通参与者的交互和自车类人安全性，模型输出需要满足预测决策规划自治性和时序稳定性，而非仅仅是无交互的开环预测；
3. AI模型的有效性难以从开环的平均指标得到证明，而闭环仿真器车辆存在与真实环境行为不同的问题，模型的有效安全性需要探索高效的证明方式；
4. 经过数据驱动后，需要AI模型能捕捉场景的差异特征，学习到场景异性的具体表现而非平均表现；
5. 需要对AI模型的输出设计有效应用和看护算法，使得自动驾驶更为安全智能。

#### 1.3 具体要求

1. 根据可参考的开源数据集，设计和实现交互预测决策规划模型，除了开环的平均指标外需要设计和给出自治指标和时序稳定指标或相关证明，需要实现闭环的评测或闭环有效性的证明，允许设计白盒算法提升AI模型输出的有效性和安全性。可在路口或人车混行等复杂场景进行通过率验证。指标参考：MinADE\_5 、 MinADE\_10 、 MissRateTopK\_2\_5 、 MissRateTopK\_2\_10 、 MinFDE\_1 、 OfRoadRate （ <https://paperswithcode.com/sota/trajectory-prediction-on-nuscenes> ） MinJointADE 、 MinJointFDE 、 MinJointMR 、 CrossCollisionRate、EgoCollisionRate、ConsistentMinJointMR（<http://challenge.interaction-dataset.com/leader-board>）
2. 可利用外部数据进行增强算法或仿真数据进行大模型预训练等（需给出所使用数据集的说明文档），并最终在测试集上进行测试，证明方法的有效性。
3. 赛题总分由竞赛得分（60%）和专家评分（40%）两部分组成。专家评分使用华为内部的中国道路的复杂场景测试数据进行指标评价和验收，考察算法的泛化性。
4. 竞赛得分部分根据选手提交的方案在测试集（包括公开测试集和非公开测试集）上的运行结果给出。
5. 专家评分由评委组对选手所提交的方案的新颖性、合理性等进行打分。因此，参赛选手还需要提交模型代码（用于非公开测试集评估）、模型说明文件（用于报告模型方案以及模型在公开测试集上的结果）。

#### 1.4 咨询专家及联系邮箱

李洪业 - [lihongye5@huawei.com](mailto:lihongye5@huawei.com)

柳肖雪 - [liuxiaoxue2@hisilicon.com](mailto:liuxiaoxue2@hisilicon.com)

## 1.5参考数据集

<https://www.nuscenes.org/nuplan#download> (https://www.nuscenes.org/nuplan#download)

<https://paperswithcode.com/sota/trajectory-prediction-on-nuscenes> (https://paperswithcode.com/sota/trajectory-prediction-on-nuscenes)

## 1.6参考文献

- [1] Interaction-Based Trajectory Prediction Over a Hybrid Traffic Graph (IROS2020)
- [2] Implicit Latent Variable Model for Scene-Consistent Motion Forecasting (ECCV2020)
- [3] Scene Transformer: A unified architecture for predicting multiple agent trajectories (ICLR2022)
- [4] Occupancy Flow Fields for Motion Forecasting in Autonomous Driving (RAL2022)
- [5] “THOMAS: Trajectory Heatmap Output With Learned Multi-Agent Sampling “, (ICLR2022)

## 2. 华为赛题二

### 2.1题目名称

适用多场景的通用化时序预测算法

### 2.2题目描述

随着云计算的规模增大、数字化程度提高，涌现出大量时间序列场景和数据，如资源的消耗序列、用户的行为序列、机房传感器的读数序列等。在针对这些数据的分析和利用中，预测问题是用户和企业最关心的任务之一，因为精准的时序预测往往是资源规划、商业运营、硬件运维等领域决策的先决条件，而这些领域的决策质量又影响了云厂商的运营效率和用户体验，定义了云厂商的核心竞争力。

云计算中的数据有着来源丰富、规模庞大、特性迥异等特点，因此传统的时序预测算法往往难以适应多种不同业务场景的需求，需要投入大量人力物力面向不同业务场景设计针对性的算法，成为了竞争力提升的一个瓶颈。将人工智能技术应用于时序预测任务是一条可行的出路，或可基于对数据特征的分析，选择合适的算法类型；或可基于神经网络等机器学习模型强大的表达能力，以单个模型实现通用化预测的目的。

主要挑战：

1) “通用性”并没有一个被广泛接受的明确定义，如果依赖具体的数据集，则一方面该数据集可能对特定类型的数据有所偏向，另一方面也会导致算法的研发倾向于过拟合该数据集。因此首要的挑战是提出并论证数据多样性和平衡性的指标，并根据该指标进行数据收集。

2) 面对拥有序列多样性的数据集，设计一个通用化的时序预测算法是一个挑战。算法可能需要处理不同长度、趋势性、周期性、采样粒度的时间序列，并且在各个序列上均给出较好的结果。

### 2.3具体要求

1) 参赛者需要定义并论证一种数据多样性和平衡性的指标，并据此基于开源数据或生成数据构建数据集，要求其中时间序列的数量不少于1000。可以根据数据特性设计一种合理的训练集和测试集的切分方式，这种切分既可以是在时间维度上的（至少保留20%作为测试集），也可以是在序列维度上的（至少保留50%作为测试集）。

2) 参赛者需要设计一种时间序列预测算法，可以是单算法也可以是集成方案，在数据集上进行训练和验证，给出算法效果和证明算法通用性的案例分析。

参赛队伍最终提交构造数据集、模型代码、方案文档和演示视频（可选），方案文档需完整描述数据集的构造方法和算法的设计思路。

构造的数据集需满足提出的数据多样性的标准，切分训练集和测试集时需至少保留20%的时间长度和50%的时间序列作为测试集。

预测算法采用MAE（Mean Average Error）和sMAPE（Symmetric Mean Absolute Percentage Error）作为评价指标，其中sMAPE指标拥有更高的优先级。在指标相近的情况下，算法的训练和推理时间将作为参考评价指标。

人工评委审核成绩从4个方面考核：结果复现评估，方案可行性分析，方案完整性评估，方案通用性评估。

### 2.4咨询专家及联系邮箱

盛镇醴 - shengzhenli@huawei.com

周乐夔 - zhoulekui@huawei.com

## 2.5 参考数据集

通用云服务：

<https://aws.amazon.com/cn/forecast/> (<https://aws.amazon.com/cn/forecast/>)

公开时序预测数据集：

<https://www.kaggle.com/datasets/yogesh94/m4-forecasting-competition-dataset> (<https://www.kaggle.com/datasets/yogesh94/m4-forecasting-competition-dataset>)

<https://robjhyndman.com/publications/the-tourism-forecasting-competition/> (<https://robjhyndman.com/publications/the-tourism-forecasting-competition/>)

<http://www.neural-forecasting-competition.com/downloads/NN5/datasets/download.htm> (<http://www.neural-forecasting-competition.com/downloads/NN5/datasets/download.htm>)

## 3. 华为赛题三

### 3.1 题目名称

通过穿戴设备监测高糖发生次数

### 3.2 题目描述

按照ADA标准，我国成年人中糖尿病前期患病率超过35%，在糖尿病前期进行控糖能够有效降低II型糖尿病和心血管疾病的风险，具有重大意义。然而，由于成本体验等原因，糖尿病前期人群中通过有创设备监测控糖水平的人很少，智能穿戴技术的巨大进步为解决糖尿病前期人群的控糖水平监测提供了前所未有的机会。

在《持续葡萄糖监测临床应用国际专家共识(2017)》中，TIR、TAR、TBR被推荐作为CGM的关键报告参数，分别反映血糖良好控制情况、高血糖和低血糖情况。《TIR国际共识》推荐大多数糖尿病患者TIR（3.9~10.0 mmol/L）控制目标为>70%、TBR（<3.9 mmol/L）<4%；TIR每增加5%，1型糖尿病（T1DM）和2型糖尿病（T2DM）患者显著临床获益；对于老年或高风险T1DM/T2DM患者，TIR控制目标降至>50%，TBR降至<1%，旨在进一步减少低血糖、同时预防严重高血糖；T1DM合并妊娠患者TIR（3.5~7.8 mmol/L）目标为>70%。对于伴有合并症及并发症的糖尿病患者，《糖尿病管理中TIR个体化控制目标：多国专家建议(2020)》强调应根据具体情况，个体化设定TIR、TAR与TBR控制目标。

数据集由16个糖前（或者接近糖前）用户组成，血糖监测设备记录用户的葡萄糖浓度（mg/dl），穿戴记录了血容量脉冲（BVP）信号、皮肤电活动（EDA）、皮肤温度和三轴加速度计。以血糖监测设备得到的血糖值是否处于正常范围（小于7.8mmol/L），把血糖值分类为正常值和异常高值（高糖）共两类。以两分类作为金标，从穿戴设备记录的数据种提取特征，训练机器学习模型，测试人群每天的高糖发生次数。

### 3.3 具体要求

1) 建议使用7.8mmol/L作为糖尿病前期人群高糖的边界，允许在此基础上有微调，例如根据数据情况可选择7.5mmol/L等。注：参数设置不合理将导致训练和测试正负样本失衡，影响最终评分。

2) 输入:某用户一天内从无创设备提供的心率（HR）、血容量脉冲（BVP）信号等数据

3) 输出: 该用户当天血糖值超过7.8mmol/L的次数。对比的金标是根据Dexcom.csv每五分钟记录的值统计此人当天实际超过7.8mmol/L的次数。

4) 打分: 提供多天的测试数据，累计算法输出的次数与金标差异，差异越小得分越高。

5) 自由选择机器学习方法训练模型（总共有16人的数据，建议在区分训练集和测试集时按照训练8人，测试8人的比例划分，可随机选8人，最终准确度需随机多次取平均，评审时需考虑模型的过拟合等问题）

6) 参赛团队可以自行选择训练样本的粒度，自由设置模型的分类能力（例如可选五分钟粒度的样本训练二分类），最终只要给出一天血糖值超过所选阈值的次数即可。

7) 最终打榜成绩（计50%分数）。线下提供测试集，测试参数团队模型的准确性。通过线下脚本计算判分取Top10进行人工评委审核，若存在并列成绩情况，可多取候选者。

8) 最终评委组通过线下自动化脚本判分，以邮件或官网形式公布最后的打榜成绩。

9) 人工评委审核成绩（计50%分数）。评审从5个方面考核：打榜结果复现评估，方案可行性分析，方案创新性评估，方案可迁移性评估，方案实用性评估；初审评委人员待定。

### 3.4 咨询专家及联系邮箱

李露平 - liluping4@huawei.com

### 3.5 参考数据集

<https://physionet.org/static/published-projects/big-ideas-glycemic-wearable/big-ideas-lab-glycemic-variability-and-wearable-device-data-1.1.0.zip> (<https://physionet.org/static/published-projects/big-ideas-glycemic-wearable/big-ideas-lab-glycemic-variability-and-wearable-device-data-1.1.0.zip>)

附件1：第五届中国研究生人工智能创新大赛初赛作品提交规范 (<https://cpipc.acge.org.cn/sysFile/downloadFile.do?fileId=8c6a06a92a26471489c034342fbdcb82>)

附件2：第五届中国研究生人工智能创新大赛项目文档模版 (<https://cpipc.acge.org.cn/sysFile/downloadFile.do?fileId=beaef69ba51b4e3b9daf2707f3dfe21c>)

中国研究生人工智能创新大赛组委会

2023年4月28日