

## 目 录

板材表面缺陷检测 .....	1
棒材组合订单高效锯切优化 .....	14
低碳钢金相图像无监督相区分割 .....	22
高炉铁次预测 .....	35
煤气发电量预测与发电优化 .....	44

# 板材表面缺陷检测

## 一、赛题背景

随着“人工智能+”战略的深入推进，智能制造已成为传统工业培育新质生产力的核心引擎。钢铁板材（如中厚板等）作为国家重大工程的基础骨干材料，其表面质量直接关系到最终产品的安全服役性能。传统的缺陷检测高度依赖人工目视或早期规则算法，存在效率低、易漏检且难以适应多变品种等瓶颈。近年来，计算机视觉技术的飞速发展为解决这一痛点提供了新路径，使质检系统逐步具备了工业慧眼的的能力。

然而，在真实的工业产线环境中，板材表面缺陷的智能感知面临着极大的技术挑战。首先，现场成像环境复杂，板材表面常伴有水渍、油污、氧化铁皮脱落等强伪影干扰，真实缺陷特征极易被掩盖。其次，缺陷形态极其复杂：不仅存在“类间相似度高”与“类内差异性大”的特征，且空间尺度跨度极大——从毫米级的微小裂纹到贯穿板面的长条划伤。此外，在实际高质量生产线中，严重缺陷往往属于罕见样本，导致训练数据呈现典型的长尾分布。这些因素对视觉模型的特征表征、抗干扰鲁棒性及跨产线泛化能力提出了严苛要求。

基于真实的产业痛点，本赛题旨在推动视觉大模型与智能检测技术在复杂工业制造场景下的前沿探索与落地。鼓励参赛队伍突破常规实验室数据集的局限，针对具有复杂背景干扰、多尺度及长尾分布特性的真实板材缺陷图像，设计高效、鲁棒的视觉感知算法。本赛题重点考察参赛者在复杂特征提取、小目标感知、模型泛化与轻量化等方面的综合水平，引导学生完成从“感知单纯数据”到“解决实际工程痛点”的认知转变。赛题成果不仅有助于深化工业视觉理论研究，更可直接赋能钢铁冶金等高端制造产线，具有重要的学术价值与现实意义。

## 二、赛题应用场景

面向复杂工业环境的板材表面缺陷智能视觉检测技术，在钢铁冶金及高端制造业的高质量发展中具有重要的落地应用价值。通过开发高鲁棒性、强泛化能力的视觉感知模型，其成果可直接赋能以下关键工业场景：

**1. 自动化产线表面质检：**在中厚板等钢铁板材的高速轧制产线上，替代传统的人工目视巡检。通过视觉算法实时精准地定位裂纹、划伤、结疤等各类表面缺陷，实现高节拍下的在线拦截与智能定级，杜绝不合格品流向船舶制造、航空航天等对材料要求极高的下游重点工程。

**2. 复杂工况下的伪影过滤与精准识别：**真实的工业现场常伴随水渍、油污、粉尘及氧化铁皮脱落等极端干扰。具备强大特征表征能力的视觉模型能够有效区分真实缺陷与生产环境带来的伪影，在水汽遮挡或背景纹理复杂等条件下，依然保持高精度的感知能力，大幅降低系统的误检率。

3. **跨产线的视觉感知与快速部署**：针对不同生产线在相机硬件、光源条件、生产品种及厚度规格上的显著差异，现有模型常因数据分布偏移（Domain Gap）而表现不佳。具备强泛化能力的视觉算法能够有效克服此类跨域难题，使得模型能够快速适配新产线、新钢种，显著降低大规模图像采集与人工标注的成本。

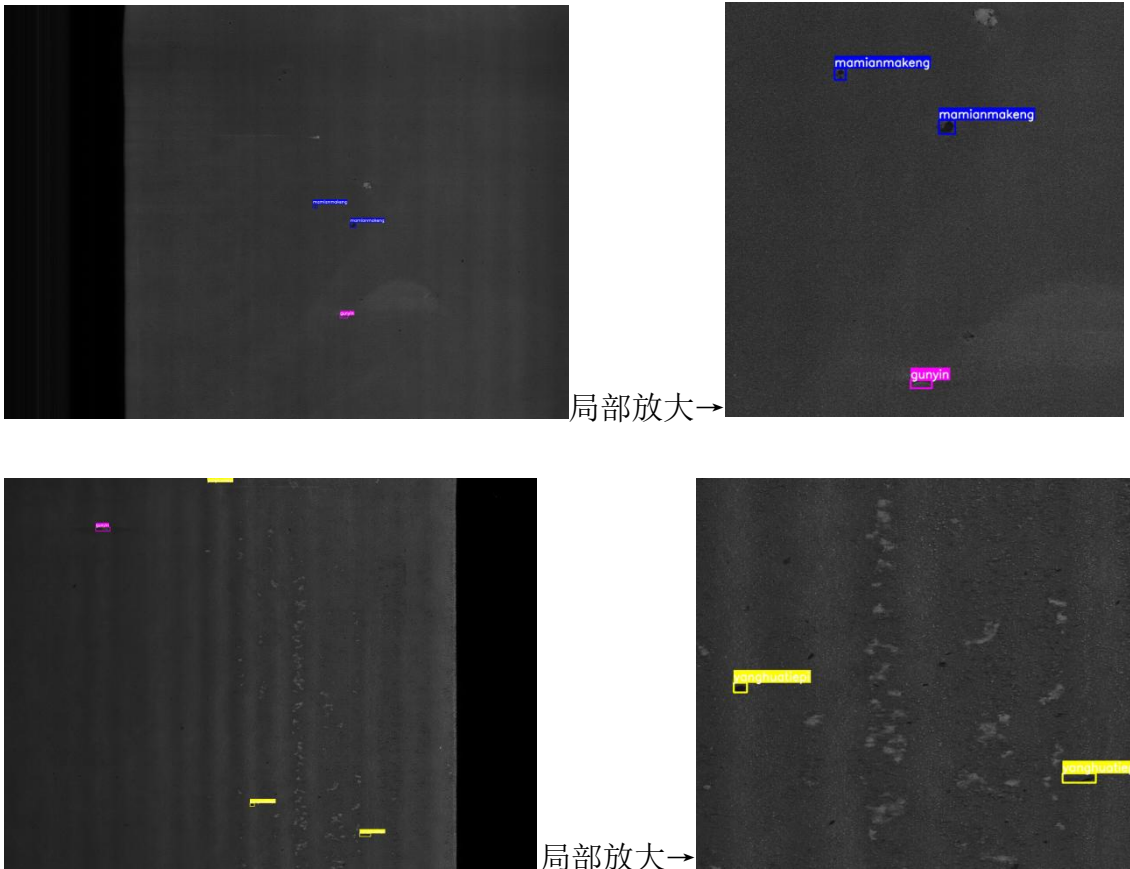
本赛题希望通过对板材表面缺陷检测算法的深入探索，推动计算机视觉与智能感知技术在真实工业制造场景中的落地应用，为改善工业质检效率、提升高端基础材料的质量管控水平提供高精度的算法支持。

### 三、赛题任务

#### （一）任务描述

参赛者需要利用赛方提供的真实工业产线高分辨率板材图像数据，掌握计算机视觉目标检测算法，能够进行复杂工业伪影（如水渍、油污、氧化铁皮）的特征过滤与抗噪处理，熟悉长尾数据分布优化及跨域泛化技术，完成以下核心任务：

1. **缺陷检测与精准定位**：构建高精度视觉目标检测模型，在强背景干扰下精准识别各类板材表面缺陷（纵裂、夹杂、麻面麻坑、氧化铁皮、辊印、异物压入、划伤、结疤、气裂），准确输出缺陷的具体类别及边界框位置。评估的核心指标为精确率（Precision）和召回率（Recall）。部分类别图例如下：



2. **极难样本与跨线泛化优化**：针对极度不平衡的长尾数据（特别是罕见但致命

的裂纹等缺陷），设计算法策略以最大化召回率，坚决杜绝漏检。

## （二）任务输入输出说明

**1. 输入说明：**图像数据主要为超高分辨率的工业相机采集图像（4096×3000 像素），图像格式为 .jpg 或 .png。标注数据（PASCAL VOC 格式）：每张图像配备一个同名的 .xml 标注文件。标注文件内详细记录了图像的基本信息与缺陷目标的精准位置。核心字段包括：

<size>：包含图像的宽（width）、高（height）和通道数（depth）。

<object>：每一个缺陷目标对应一个 <object> 模块。

<name>：缺陷类别的拼音或英文缩写（例如：zonglie 代表纵裂，jiaza 代表夹杂等）。

<bndbox>：缺陷的边界框坐标，由左上角坐标（xmin, ymin）与右下角坐标（xmax, ymax）组成，坐标值为绝对像素位置。

**2. 输出说明：**模型需对测试集中的每一张图像进行处理。参赛队伍需将所有图像的预测结果汇总，统一打包为一个 JSON 格式文件（例如 submission.json）进行提交。

系统将自动解析该 JSON 文件，并与隐藏的真实标签（Ground Truth）计算 mAP（平均精度均值）等评测指标。

提交的 JSON 文件应为一个列表（List），列表中的每个字典（Dictionary）代表一个预测出的缺陷目标。具体字段定义如下：

JSON

```
[
  {
    "image_id": "0001400166-Raw00-f_00006.jpg",
    "category_name": "gunyin",
    "bbox": [2560, 1520, 2615, 1539],
    "score": 0.95
  },
  {
    "image_id": "0001469760_Raw13_f_00003.jpg",
    "category_name": "huashang",
    "bbox": [648, 702, 1681, 750],
    "score": 0.88
  },
]
```

```
{
  "image_id": "0001388270-Raw01-f_00001.jpg",
  "category_name": "jiaza",
  "bbox": [1148, 2178, 1191, 2335],
  "score": 0.92
},
{
  "image_id": "0001400264-Raw03-f_00009.jpg",
  "category_name": "jieba",
  "bbox": [1108, 2, 1213, 118],
  "score": 0.91
},
{
  "image_id": "0001388609-Raw00-f_00007.jpg",
  "category_name": "mamianmakeng",
  "bbox": [1459, 1524, 1502, 1561],
  "score": 1.0
},
{
  "image_id": "0001417541-Raw00-f_00005.jpg",
  "category_name": "qilie",
  "bbox": [1084, 2243, 1150, 2400],
  "score": 1.0
},
{
  "image_id": "0001389511-Raw02-f_00002.jpg",
  "category_name": "yanghuatiepi",
  "bbox": [2856, 1290, 2874, 1333],
  "score": 1.0
},
{
```

```

    "image_id": "0001400166-Raw01-f_00001.jpg",
    "category_name": "yiwuyaru",
    "bbox": [3648, 2607, 3688, 2651],
    "score": 1.0
  },
  {
    "image_id": "0001388270-Raw01-f_00001.jpg",
    "category_name": "zonglie",
    "bbox": [3153, 1473, 3184, 2037],
    "score": 1.0
  }
]

```

#### 四、数据集及数据说明

##### （一）数据来源

本赛题提供的数据集为自行采集的真实工业产线数据。所有图像数据均来源于南京钢铁股份有限公司（中厚板等核心产线）的表面质量在线检测系统，由高分辨率工业线阵相机在高速轧制工况下实时抓拍。经过严格的人工筛选与高精度专家标注，构建了本赛题的专属数据集。

所有数据确保来源合法合规，并明确仅授权用于本次竞赛的学术探索、算法验证及非商业用途。

##### （二）数据规模与划分

赛题所提供的数据集共包含约 5000 张高分辨率工业图像。为全面、客观地评估参赛模型的性能，数据集将分为训练集、初赛测试集、复赛测试集、半决赛测试集，各阶段数据划分如下：

训练集：包含约 3200 张带有精确人工标注框的缺陷图像，用于参赛队伍训练视觉检测模型。该训练集包括 9 类，由于类间差异较小，训练集已经类别合并，包括：

合并后类别	原始类别	数量
结疤	结疤、重皮、轧损	604
纵向裂纹（纵裂）	纵裂	296
气泡状裂纹（气裂）	气裂、横裂	28
夹杂	夹杂	113
异物压入	掉渣、边丝压入、压坑、异物压入、双	464

	边剪压痕、定尺剪压痕	
划伤	划伤	60
麻面麻坑	麻面、麻坑	309
氧化铁皮	——	358
辊印	鼓包、凸块	236

**初赛测试集：**包含约 800 张无标签图像，数据分布与训练集基本一致，用于初赛阶段的客观评测。

**复赛、半决赛测试集：**包含约 1000 张无标签图像。此部分数据引入了不同产线的复杂样本，用以考察模型在分布偏移（Domain Shift）下的泛化能力。

### （三）数据特点与预处理说明

本数据集高度还原了真实恶劣的工业感知环境，具备以下显著特征：

1. 数据类型与尺度：数据为超高分辨率的单通道灰度图（典型分辨率为 4096×3000）。目标缺陷的尺度跨度极大，既有横跨画面的长条状“纵裂（zonglie）”，也有仅占数十像素的细小“夹杂（jiaza）”。

2. 长尾分布（Long-tail）：由于实际高质量产线中严重缺陷极少，数据集的类别呈现典型的长尾分布。普通轻微缺陷样本较多，而致命性缺陷样本稀缺。

3. 复杂背景干扰：图像背景中自然叠加了真实产线的水渍、油污、粉尘及正常水切痕迹等高强度伪影噪声。

4. 预处理说明：为了降低参赛队伍的数据清洗成本并聚焦于算法本身，赛方已对原始采集数据进行了基础的预处理：

5. 数据清洗：剔除了由于相机光源频闪导致的全黑/全白失效帧，以及完全无意义的空跑背景帧。

6. 格式标准化：将不同底层传感器采集的 RAW 数据统一解码、对比度拉伸并转换为标准的 8-bit .jpg 格式图像。

7. 未做裁切与归一化：为保留缺陷的全局上下文信息，赛方未对原始超大分辨率图像进行强制切图或缩放归一化。参赛队伍需根据自身采用的基础大模型或目标检测网络（如 YOLO 系列等），自行设计数据加载时的缩放（Resize）、切图（Crop）及张量归一化（Normalization）策略。

### （四）数据格式

赛题训练数据集采用经典的 PASCAL VOC 格式进行组织。每一组训练数据由一张图像文件和对应的同名文本标注文件构成。

图像文件：标准 .jpg 格式，文件命名规则如 0001388270-Raw01-f\_00001.jpg。

标注文件：标准 .xml 格式。文件中包含了图像的基础属性（宽 width、高

height、通道数 depth) 以及缺陷目标的详细信息。

针对图像中的每一个缺陷，标注信息存储在 <object> 标签内，核心字段包括：

<name>：缺陷类别的唯一标识（如 zonglie、jiaza 等）。

<bndbox>：目标缺陷的绝对像素坐标边界框，由左上角 [xmin, ymin] 和右下角 [xmax, ymax] 的绝对相对数值（整数）构成，取值范围为 [0, 图像宽/高]。

## 五、算法设计要求

### （一）模型类型

鼓励参赛者采用并深度优化适配工业视觉场景的深度学习目标检测算法。优先支持以主流计算机视觉架构为核心的技术路线，包括但不限于：基于卷积神经网络（如 YOLO 系列）的高效检测算法、基于视觉 Transformer（如 DETR 系列）的检测模型等，充分发挥深度学习算法在细粒度特征提取、复杂背景抗干扰以及工业级高节拍实时推理等方面的核心优势。

### （二）创新性

鼓励参赛队伍围绕真实工业环境下的视觉检测痛点提出创新性改进方案，重点提升模型对多尺度缺陷的精准感知与跨产线泛化能力。例如，针对板材表面强伪影干扰（如水渍、油污掩盖）或致命缺陷（如微细裂纹）漏检等短板，设计专属的抗噪特征提取网络、多源特征融合机制或跨域适应（Domain Adaptation）模块，强化模型在长尾数据分布下的识别精度与定位准确性。

## 六、性能指标要求

本赛题聚焦于板材表面缺陷的精准检测。针对真实工业产线“宁可少量误报，绝不可漏检”的实际业务需求，本赛题采用精确率（Precision）与召回率（Recall）作为评估参赛模型性能的核心指标。

两项核心指标的计算公式为：

$$\text{Precision} = \frac{TP}{TP+FP}$$

$$\text{Recall} = \frac{TP}{TP+FN}$$

参赛者的算法模型必须在保证基础精确率的前提下，尽可能追求高召回率。

## 七、功能要求

参赛者提交的解决方案应具备以下核心功能并满足相应的验收标准：

1. 精准检测与高召回：算法能够在复杂工业背景（如存在水渍、油污等伪影干扰）下，准确识别并定位板材表面缺陷。针对裂纹等致命缺陷，必须具备极高的召回能力（防漏检）。

2. 跨场景鲁棒性：面对未见过的跨产线、不同规格的测试数据以及长尾分布样本时，模型不应出现性能的大幅波动，需保持优秀的泛化与适应能力。

3. 规范化结果生成：模型需能独立、全自动地在官方测试集上完成端到端推理，并严格按照赛题指定的格式（如标准的 JSON 文件）输出预测结果。

4. 功能测试与验收：赛事方将在统一的标准 GPU 算力环境下对提交方案进行自动化黑盒测试。测试用例将涵盖常规缺陷、强干扰伪影。参赛代码须在规定时间内无报错跑通，且核心指标（精确率与召回率）需达到专家组设定的基准线。

## 八、开发环境

### （一）软件环境

本赛题对参赛队伍的编程语言和具体框架不做强制性限制，但为保证开发效率与最终方案的顺利复现，官方提供以下环境建议：

1. 编程语言：推荐使用 Python 3.8 及以上版本

2. 深度学习框架：推荐使用 PyTorch（建议 2.0 及以上版本）或 TensorFlow（建议 2.10 及以上版本）等主流深度学习框架。允许使用开源的预训练权重与计算机视觉模型库（如 torchvision、MMDetection、YOLOv8/v10 官方库等）进行算法搭建。

3. 数据与图像处理库：鉴于本赛题包含超高分辨率的工业视觉数据，推荐使用具备高效处理能力的库，如 OpenCV-Python、Pillow 用于图像的读取与增强变换；NumPy、Pandas 用于标签数据的解析与矩阵运算；Matplotlib 用于可视化分析。

4. 代码复现要求：无论使用何种开源框架，参赛者需确保最终提交的代码具备完整的环境依赖说明（如提供 requirements.txt），保证赛事方能够无缝复现模型训练与推理过程。

### （二）硬件环境

1. 硬件限制：本赛题对参赛队伍使用的 CPU 型号、内存大小、GPU 型号等硬件资源不做强制限制，赛事方亦不统一提供算力支持。

2. 硬件建议：考虑到高分辨率工业图像的加载以及目标检测模型的显存占用，建议使用具备较大显存（如 16GB 及以上）的 GPU 设备开展模型训练与推理，以保障模型训练的效率和稳定性。

3. 云端算力平台：参赛者可根据自身模型规模与硬件条件，灵活选择本地工作站或借助云端计算平台进行开发。允许并鼓励使用公有云 AI 训练平台，如阿里云天池、百度 AI Studio、腾讯云 TI-ONE、火山引擎方舟平台等获取计算资源。

## 九、成绩评价

### （一）输出数据格式要求

为保证自动化评测的顺利进行，参赛者必须确保提交的预测结果完全符合官方评测脚本的解析标准。若因格式不规整导致评测程序匹配失败，最终提交结果将无法得分。

参赛者需对所有测试图像进行目标检测识别，并将全量预测结果统一写入并导出一个标准的 JSON 文件。每个缺陷目标的预测字段必须严格包含以下内容：

图像的唯一标识（如 `image_id`）

预测的缺陷类别标签（如 `zonglie`、`jiaza` 等）

目标的矩形边界框坐标 `[xmin, ymin, xmax, ymax]`（绝对像素值，且需满足  $xmin < xmax, ymin < ymax$ ）

预测置信度得分（0 到 1 之间的浮点数）

## （二）评价阶段与成绩构成

本赛题的成绩评价贯穿初赛、复赛、半决赛和总决赛四个阶段。为了全面衡量算法在复杂工业缺陷场景下的稳定表现，避免过度依赖单一阶段的刷榜结果，赛事方将采用以下成绩计算体系：

1. 初赛阶段：主要用于参赛者熟悉板材表检任务、验证基础视觉检测方案与完成晋级。初赛的客观评测得分不计入最终线上综合成绩。

2. 复赛与半决赛阶段：将引入长尾分布更显著、跨产线伪影干扰更强的测试集。复赛与半决赛的客观评测成绩将共同构成“线上综合成绩”。其中，复赛线上成绩占 40%，半决赛成绩占 60%。

## 十、解题思路

### （一）数据预处理策略

针对工业相机采集的超高分辨率图像以及表面的水油渍干扰，首要任务是设计合理的图像缩放或局部裁切策略，以防微小缺陷特征在输入模型前丢失；同时，需引入针对性的数据增强手段，提升模型在肮脏背景下的抗干扰能力。

### （二）基础检测模型适配

选用成熟的视觉目标检测网络作为基座。考虑到板材缺陷既有贯穿全图的长条划伤，也有极其细小的夹杂，模型设计的重点在于强化“多尺度特征提取与融合”，确保不同大小的缺陷都能被有效捕捉。

### （三）核心难点攻坚：长尾与泛化

这是决定最终成绩上限的关键环节。参赛者需重点突破两大工程难题：

1. 防漏检（长尾分布）：致命缺陷样本通常极少，需通过样本重采样或优化损失函数等策略，强迫模型对稀有致命缺陷保持高度敏感。

2. 抗偏移（跨线泛化）：针对复赛中可能出现的新产线、新规格数据，需探索有效的特征对齐或泛化方法，避免模型在面对未见过的数据时出现性能骤降。

### （四）推理与后处理优化

在生成最终预测结果时，需设计合理的边界框过滤机制以剔除重复或低置信度的误报框。最后，将局部预测结果精准映射回原始大图的绝对坐标系中，并严格按照官

方要求的格式生成提交文件。

## 十一、赛题约束条件

### (一) 算法约束

禁止调用商业闭源 API：严禁调用任何商业闭源视觉大模型或在线推理 API（如 GPT-4V、Gemini API 等商业接口）。参赛队伍仅允许使用开源视觉目标检测模型或自研算法框架进行自主开发、微调与本地部署。

1. 禁止人工干预测试结果：禁止采用手工标注、人工复核等方式修改测试集结果。所有预测的类别标签与边界框必须完全由算法模型自主推理输出，严禁任何形式的人工干预。

2. 禁止多模型无脑集成：为贴合真实工业产线对检测算法高实时性、低显存占用的严苛要求，最终提交方案禁止将多个不同网络结构或不同训练阶段的模型进行简单粗暴的集成（如多模型投票法、加权平均法等）。

### (二) 数据使用约束

1. 允许合规扩充训练数据：为鼓励算法创新并挑战更高精度，参赛队伍在训练阶段允许自行扩充数据集（例如通过数据合成、自行采集或引入开源的外部缺陷数据集参与训练）。但参赛者必须在最终提交的技术报告中，如实、详细地说明所扩充数据的来源、规模及合成方式。

2. 严格的数据保密与防泄露红线：赛事官方提供的所有中厚板产线图像数据及专家标注文件均属企业生产机密。严禁以任何形式将官方数据集泄露、外传（包括但不限于公开分享、上传至 GitHub/网盘等公开代码仓、转售或用于非本次赛事相关的学术发表及商业用途）。参赛队伍需妥善保管下载的数据，一经发现泄露行为，赛事方将立即取消其参赛资格，并保留追究相关法律责任的权利。

3. 测试集隔离约束：严禁私自对官方测试集进行人工标注或篡改，严禁将测试集数据（含打伪标签）以任何形式混入模型训练环节中。

## 十二、参考资源

### (一) 文献资料

1. 黄远民, 易铭, 杨曼, 等. 基于机器视觉的板材表面缺陷自动检测方法研究[J]. *Artificial Intelligence and Robotics Research*, 2019, 8: 109.

2. Chang Z, Cao J, Zhang Y. A novel image segmentation approach for wood plate surface defect classification through convex optimization[J]. *Journal of Forestry Research*, 2018, 29(6): 1789-1795.

3. Hu L, Zhou M, Xiang F, et al. Modeling and recognition of steel-plate surface defects based on a new backward boosting algorithm[J]. *The International Journal of Advanced Manufacturing Technology*, 2018, 94(9): 4317-4328.

## （二）在线课程

1. 【2025 新版【机器学习十大算法全集】17 分钟让你看懂所有机器学习算法！一口学透回归算法、聚类算法、决策树、随机森林、神经网络、贝叶斯算法、支持向量机、神经网络等】 [https://www.bilibili.com/video/BV1BoNvzfELN/?share\\_source=copy\\_web&vd\\_source=8eeba1d0160d38791a8a4db667ba5c84](https://www.bilibili.com/video/BV1BoNvzfELN/?share_source=copy_web&vd_source=8eeba1d0160d38791a8a4db667ba5c84)

2. 【机器学习深度学习项目实战（波士顿房价预测、图像识别水果分类、新闻分类）】 [https://www.bilibili.com/video/BV1Af4y1x7js/?share\\_source=copy\\_web&vd\\_source=8eeba1d0160d38791a8a4db667ba5c84](https://www.bilibili.com/video/BV1Af4y1x7js/?share_source=copy_web&vd_source=8eeba1d0160d38791a8a4db667ba5c84)

## 十三、提交要求

### （一）初赛、复赛与半决赛（线上评测阶段）提交要求

在初赛、复赛及半决赛的线上评测阶段，参赛者需在官方提供的各阶段对应测试集上进行模型推理，并在规定时间内提交预测结果文件。

提交文件格式统一为 JSON 格式（如命名为 `submission.json`）。

提交内容需严格按照前述“数据格式要求”，包含测试图像文件名、预测的缺陷类别名称、边界框坐标（BBox）以及置信度得分。

线上各阶段均以该 JSON 结果文件作为评测平台自动打分、排行榜展示及晋级评定的唯一依据。

提交限制：每天最多提交 5 次，排行榜实时更新，以最后一次提交为准。

### （二）半决赛、总决赛（线下答辩阶段）提交要求

半决赛、总决赛的参赛团队，除了需提交最终的测试集预测结果外，还必须按赛事组委会要求提交完整的工程与技术文件，以供专家评审组进行代码复核与答辩评审。具体提交清单如下：

1. 完整的算法代码与运行说明：提交包含数据预处理、模型构建、训练、验证、预测推理以及结果后处理等各个环节的完整 Python 代码。代码需具备良好的工程结构与清晰的规范化注释。提供详尽的运行说明文档（如 `README.md`）及完整的环境配置文件（如 `requirements.txt`），确保赛事方评审人员能够依据文档顺利复现出与线上榜单一致的实验结果。

2. 可执行的模型文件：提交对应最终最佳成绩的预训练权重与最终训练完毕的模型权重文件（如 `.pt`、`.pth` 等），并附带模型加载与使用的具体说明。

3. 技术方案报告（PDF 格式）：提交内容详实的技术报告，详细阐述整体算法框架与创新点。报告内容应重点包含：算法设计动机、数据增强方法、模型网络架构、针对板材表检核心难点（如长尾分布缺陷的召回优化、强伪影干扰的过滤、跨产线数据的泛化策略等）的定制化设计、关键模块消融实验分析等。

4. 作品海报：尺寸为 A1（841mmx594mm），电子版分辨率为 300dpi，文件统一保存为 PNG 格式。

5. 答辩演示文稿（PPT）：提交用于总决赛现场线下展示与专家答辩的 PPT 文件，以及赛事组委会要求的其他补充证明材料。

#### 十四、奖励设置

为了鼓励参赛选手参赛积极性，激发在复杂制造场景下的技术创新潜能，本赛题不仅设立了丰厚的现金奖励，更针对优秀的青年算法人才开辟了专属的职业发展绿色通道。具体激励设置如下：

##### （一）赛事奖金设置

本赛题将根据总决赛的最终综合成绩（含线上客观评测与线下答辩），对全国总决赛一等奖前六名参赛团队按照如下标准颁发奖金：

1. 冠军奖：第 1 名，奖金 8000 元/每团队
2. 亚军奖：第 2-3 名，奖金 6500 元/每团队
3. 季军奖：第 4-6 名，奖金 3000 元/每团队

##### （二）人才招聘专属特权

为加速“人工智能+”创新成果向新质生产力的转化，切实打通产学研用链路，针对在本赛题中展现出卓越代码能力与工程落地思维的优秀参赛选手（不限于前六名获奖团队成员），出题方还将额外提供以下极具含金量的职业发展激励：

##### 1. 核心岗位面试直通车：

在复赛或总决赛中表现突出的核心算法选手，将获得人工智能研究院算法研发岗位的“校招面试直通车”。凭此卡可直接免除在线笔试与简历初筛环节，直通终审技术面，优先锁定高潜算法人才核心席位。

##### 2. 专项精英实习计划：

定向开放工业计算机视觉、大模型前沿应用等核心研发方向的优质实习岗位。受邀实习生将直接参与真实的高级别产线视觉项目，享有充足的算力资源支持与资深算法研究员的 1v1 业务指导，积累宝贵的工业界实战经验，实习期间表现优异者可获全职 Offer 提前转正机会。

#### 十五、其他说明

**公平性：**严禁任何形式的作弊行为，包括但不限于数据泄露、模型预训练数据与测试数据重叠、抄袭他人代码等。一经发现，立即取消参赛资格，并追究相关责任。

**知识产权：**参赛者提交的作品必须为原创，未在其他比赛中获奖或公开发表。比赛主办方有权对参赛作品进行展示、宣传等相关活动，但知识产权仍归参赛者所有。

#### 十六、联系方式

赛题交流 QQ 群：333650987

邮 箱: wujialin@njsteel.com.cn

报名官网: [www.aicomp.cn](http://www.aicomp.cn)

## 棒材组合订单高效锯切优化

### 一、赛题背景

棒材锯切是钢铁生产线成品交付的核心工序，棒材是由加热钢坯通过多道轧辊从立方体变成圆柱体，类似“拉面条”把一块“方面团”拉成“超长细面条”。因为现场设备的限制，“超长细面条”无法直接上锯切辊道，需要将其分段，也就是把“细面条”分成多段“细面条”一起进行批量锯切，锯切方案设计（同时对几根“细面条”进行锯切，每根“细面条”有多长等等）直接决定产线产能释放与综合成材率。并且随着钢铁行业向多品种、小批量柔性生产转型，订单个性化、定制化趋势显著，大量非经济订单（倍尺长度、重量与锯切节奏不匹配）无法铺满锯切辊道，导致锯切产能闲置、设备利用率低下。

当前棒材锯切排产仍普遍依赖人工经验与固定规则，存在两大难以突破的核心痛点：其一，人工规则基于简化的经验公式，难以精准平衡倍尺长度、设备动作周期与生产节奏的复杂耦合关系，面对指数级增长的订单组合空间，无法找到全局最优解，导致小批量低效锯切频发，锯切刀数冗余度高；其二，订单繁杂度呈指数级上升，人工无法对海量同钢种、同规格订单进行系统性聚类与跨订单整合，只能按订单顺序依次排产，推高了锯切批次数量与单位产品生产成本。同时，不同操作人员的技能水平与操作习惯差异显著，导致排产结果波动大，产线运行稳定性不足。

然而，仅依靠简单的订单聚类与规则匹配远远不够，在实际生产中，棒材锯切排产需同时兼顾多重刚性约束：一是冷床承载约束，组合订单总长度与总重量必须严格控制在冷床上下限范围内；二是锯切工艺约束，倍尺长度必须为订单定尺长度的整数倍且预留合理加工余量。

本赛题旨在借助先进组合优化算法，突破传统人工排产的技术瓶颈，推动棒材锯切排产向数字化、智能化方向升级。参赛者需利用订单数据与设备工艺参数，构建高效的全局寻优模型，在满足冷床承载、锯切工艺等多重约束的前提下，设计最优的订单组合策略与锯切参数方案，明确各批次订单组成、倍尺长度设定与锯切批次划分，最终实现成材率最优、锯切刀数最少化与产线产能最大化。

### 二、赛题应用场景

以钢铁企业棒材生产线轧制后、锯切前订单排产为典型应用场景。该场景下，企业需从海量未排产订单池中筛选同钢种、同规格订单进行组合，使组合后的订单总长度适配冷床承载能力，且通过尽可能铺满锯切辊道实现高效批量锯切。

现场生产中，不同订单的定尺长度、重量需求差异显著，人工需逐一计算订单倍尺长度并尝试组合，不仅耗时费力，且极易出现组合不合理导致的锯切刀数过多、冷床空间浪费等问题。开发能够自动完成订单聚类、组合寻优与锯切方案生成的智能系

统，已成为棒材产线降本增效的共性技术需求。

### 三、赛题任务

参赛团队需根据订单、钢坯等输入，构建棒材组合订单经济锯切优化模型，实现同钢种、同规格订单的智能组合与锯切参数自动优化。系统需重点攻克大规模订单组合空间的高效寻优难题，在严格满足工艺与设备约束的前提下，输出最优订单匹配组合方案及相应的锯切参数设定。具体任务包括：

1. 开发订单数据预处理与校验模块，实现同钢种、同规格订单的自动聚类与异常数据过滤；
2. 设计高效的组合优化算法，在满足所有约束条件的前提下，达到组合订单总锯切刀数最小化、成材率最大化、组合订单覆盖率最大化、求解时间最小化的综合最优；
3. 生成可直接指导生产的锯切方案，包含组合订单编号、各订单倍尺长度、锯切批次划分等关键信息。

### 四、数据集及数据说明

#### 模型输入：

##### （一）订单数据：

文件名称	主要说明
order_quarter.csv 初赛：5k 条订单数据 复赛：1w 条订单数据 决赛：2w 条订单数据	订单号，订单钢种，订单重量，订单规格（直径和定尺长度），密度

##### （二）钢坯数据：

文件名称	主要说明
blank_used.csv 初赛：2 种钢坯规格 复赛、决赛：5 种钢坯规格	钢坯规格（宽度、厚度和长度），密度

#### 模型约束：

##### （三）约束条件：

文件名称	主要说明
constraint.txt	覆盖同质性，冷床长度等约束条件

#### 模型输出：

##### （四）输出结果：

结果为 list<json>形式，以\*.json 文件格式保存并上传。示例：

```
[
  {
    "orders": [
      "A20260103"
    ],
    "length_scheme": [
      {
        "A20260103": 83.2
      }
    ],
    "counts": [
      71
    ],
    "blank_type": 1,
    "blank_counts": [
      4
    ]
  },
  {
    "orders": [
      "A20260104",
      "A20260105"
    ],
    "length_scheme": [
      {
        "A20260105": 109.25,
        "A20260104": 25.6
      },
      {
        "A20260104": 70.4
      }
    ],
    "counts": [
      71,
      71
    ],
    "blank_type": 1,
    "blank_counts": [
      7,
      4
    ]
  }
]
```

list 中的每个 json 为一个订单组合方案。示例结果提供了 3 个订单的可行组合与

锯切方案，第一个组合方案为订单单独锯切，第二个组合方案为 2 个订单组合锯切。参赛选手上传的解提供的锯切方案应覆盖所有订单，并平衡成材率与组合订单覆盖率以提高得分。每个订单组合方案的详细参数介绍如下表所示：

字段与格式	主要说明
orders: list<string>	[订单编号 1, 订单编号 2, ...] 表示该方案组合了哪些订单。 例如：["A20260104", "A20260105"]表示编号为 A20260104, A20260105 的订单放在一起进行组合锯切，["A20260103"]表示编号为 A20260103 的订单单独锯切。
length_scheme: list<json> 单位：m	[[{订单编号 1: 分配长度 1, 订单编号 2: 分配长度 2, ...}, ...] 表示锯切方案，每个 json 表示每次冷床上的具体长度分配方案。 例如： [{"A20260105": 109.25, "A20260104": 25.6}, {"A20260104": 70.4}]表示本锯切方案下，第一轮上冷床的棒材有 109.25m 的长度用于完成订单 A20260105，25.6m 的长度用于完成订单 A20260104；第二轮上冷床的棒材有 70.4m 的长度用于完成订单 A20260104。
counts: list<int>	[辊道并列支数 1, 辊道并列支数 2, ...] 表示辊道上横向并列摆放的棒材支数，与锯切方案一一对应。 例如：[71, 65]表示本锯切方案下，第一轮在冷床辊道上横向并列摆放了 71 支棒材，第二轮在冷床辊道上横向并列摆放了 65 支棒材。
blank_type: int	钢坯类型 id 表示锯切方案使用的钢坯。 例如：1，表示本锯切方案使用 blank_used.csv 中的第一种钢坯。
blank_counts: list<int>	[钢坯数量 1, 钢坯数量 2, ...] 表示每轮上冷床消耗的钢坯数量，与锯切方案一一对应。 例如：[7, 4]，表示第一轮上冷床需要使用 7 支钢坯，第二轮上冷床需要使用 4 支钢坯。

## 五、算法设计要求

(一) 鼓励参赛者结合启发式策略提升大规模组合空间的寻优效率。可尝试融合

遗传算法、模拟退火、禁忌搜索等智能优化算法的优势，进一步提升求解质量。

(二) 算法需具备良好的泛化能力，能够动态适应订单规模、订单结构的波动，在订单数量从几十到几万的范围内均能保持稳定的求解性能。

(三) 算法设计应兼顾实时性与可靠性，单批次订单组合的求解时间应不大于 10 秒，能够满足工业现场排产的实时性要求。

(四) 建议核心算法模块由参赛队自主设计或基于开源框架进行合理扩展，若使用商业求解器，则仅限使用杉数求解器 COPT。

## 六、性能指标要求

本赛题以实际生产订单数据为基准测试集，构建标准化的性能评估体系，通过离线仿真评估算法性能。性能评价主要包括以下三个方面：

### (一) 总锯切刀数

计算方法：总锯切刀数 = 定尺长度总锯切次数 + 切头切尾各一次

例如：棒材总长 50 m，定尺长度为 3 m。考虑两端切损共 2 m（每端 1 m），有效长度为 48 m，可切出 16 根定尺棒材，无剩余。锯切次数为：两端切头/尾 2 次 + 切分 16 根所需的 15 次 = 17 次。

### (二) 成材率

计算方法：成材率 = 成材（定尺）总重量/钢坯总重量

(三) 组合订单覆盖率：组合锯切的订单数量占总订单数量的比例，仅统计**多订单组合方案**中的订单。

计算方法：组合覆盖率 = 组合订单总数量 / 测试集订单总数量 × 100%

### (四) 平均求解时间：单批次订单组合的平均求解时间（初赛不涉及）

## 七、功能要求

本系统要求利用订单数据与设备工艺参数，构建高效的全局寻优模型，输出结果包括组合方案（哪些订单放在一起锯切）、锯切方案（几个锯切批次，每个锯切批次如何分配订单）和钢坯方案（使用哪种钢坯，使用几根）。

## 八、开发环境

### (一) 软件环境

1. 参赛者开发过程的软件系统不限（Windows、Linux、Mac 等），最终代码应当保证在 Linux 系统下测试通过。开发工具（IDE）不限。

2. 开发语言应当以 Python 为主语言，核心控制算法模块可使用 C++ 进行性能优化。若采用其他语言，参赛者应当合理封装，并以 Python 作为胶水语言实现相关接口。

### (二) 硬件环境

硬件开发环境不限。最终代码应当至少在标准机器（32 核心 CPU，32GB 内存，

显存 24G) 硬件资源下实现稳定运行。

## 九、成绩评价

**1. 指标权重：**总锯切刀数占 40%、成材率占 30%，组合订单覆盖率占 20%，平均求解时间占 10%（初赛不作为评价指标）

总得分 =  $0.4 * (\text{基准锯切刀数} / \text{选手锯切总刀数}) + 0.3 * \text{成材率} + 0.2 * \text{组合订单覆盖率} + 0.1 * (\text{基准求解时间} / \text{选手求解时间})$

### 2. 其他评分因素：

**a. 代码规范性与可部署性（占附加分的 50%）：**代码结构模块化、注释完整、可直接部署于工业现场，无平台依赖，得满分；耦合度高、可移植性差或存在明显缺陷，酌情扣分。

**b. 技术报告与答辩表现（占附加分的 50%）：**技术方案完整、逻辑清晰，包含算法设计、仿真验证、创新点说明及工程价值分析，答辩能准确回应技术问题，得满分；报告内容不完整、逻辑混乱或答辩表现差，酌情扣分。

## 十、解题思路

构建基于组合优化的棒材组合订单经济锯切排产系统，以总锯切刀数最少化、成材率最大化为目标，综合考虑冷床承载约束、锯切工艺约束与生产节奏约束，自动生成最优的订单组合方案与锯切参数设定。针对大规模订单组合空间的寻优难题，融合遗传算法、模拟退火等元启发式算法平衡求解质量与速度。同时增强算法的鲁棒性与工程实用性。

并非所有订单都需要进行组合锯切，锯切方案中也可给出单订单进行锯切的锯切方案。赛题目标包括锯切刀数、成材率、组合覆盖率与求解时间，不同的目标具有不同权重，选手应在解满足约束的基础上尽可能达到解的综合最优。

## 十一、赛题约束条件

### （一）算法约束

允许使用开源库与框架：NumPy/Pandas（数据处理）、Matplotlib（数据可视化）、Geatpy（开源群智能优化算法）等等，禁止使用封装式 API 直接输出结果。

### （二）求解器使用约束

允许使用自研求解器或开源求解器。评分环境仅具备 COPT 的商业授权，若使用商业求解器，则必须使用 COPT，不得使用其他商业求解器或任何破解版商业求解器。

### （三）数据使用约束

1. 数据来源限定：仅允许使用赛方提供的，严禁使用其他公开 / 私有数据集。
2. 数据保密要求：赛事提供的所有数据集仅可用于本次赛事相关研究与参赛，禁止泄露、传播或用于非赛事相关的商业用途，违反者取消参赛资格。

## 十二、参考资源

1. 张峰, 魏洪云, 刘浩, 等。一种棒材锯切组合模型: CN113238529B [P]. 2022-09-13.

2. 组合优化相关技术

## 十三、提交要求

### (一) 初赛提交内容

1. 结果文件: 结果以 json 格式保存, 按“初赛结果\_团队名.json”命名, 压缩为“初赛结果\_团队名.zip”(压缩包内无嵌套文件夹);

2. 技术摘要: PDF 格式 ( $\leq 2$  页), 简述算法思路、核心逻辑;

3. 提交限制: 每天最多提交 5 次, 排行榜实时更新, 以最后一次提交为准。

### (二) 复赛提交内容

1. 结果文件: 结果以 json 格式保存, 按“复赛结果\_团队名.json”命名, 压缩为“复赛结果\_团队名.zip”(压缩包内无嵌套文件夹);

2. 完整代码: 含模型脚本, 需提供详细注释及运行说明 (README.md);

3. 技术报告: PDF 格式 ( $\geq 3000$  字), 详述算法架构、创新点等情况;

4. 算法实现: 提交求解算法的完整代码 (需支持完整复现最终求解结果);

5. 提交限制: 每天最多提交 5 次, 以最后一次提交为准。

### (三) 半决赛提交内容

1. 结果文件: 结果以 json 格式保存, 按“复赛结果\_团队名.json”命名, 压缩为“半决赛结果\_团队名.zip”(压缩包内无嵌套文件夹);

2. 完整代码: 含模型脚本, 需提供详细注释及运行说明 (README.md);

3. 技术报告: PDF 格式 ( $\geq 3000$  字), 详述算法架构、创新点等情况;

4. 算法实现: 提交求解算法的完整代码 (需支持完整复现最终求解结果);

5. 提交限制: 复赛期间共 5 次提交机会, 以最后一次提交为准。

### (四) 决赛提交内容

1. 全流程工程包: 含代码、模型、一键运行脚本;

2. 技术报告升级版: 补充决赛测试集实验分析、泛化能力验证;

3. 演示视频:  $\leq 10$  分钟, 展示算法运行流程与求解结果;

4. 答辩 PPT:  $\leq 20$  页, 提炼核心技术与成果;

5. 作品海报: 尺寸为 A1 (841mmx594mm), 电子版分辨率为 300dpi, 文件统一保存为 PNG 格式。

## 十四、奖金设置

为了鼓励参赛选手参赛积极性, 激发在复杂制造场景下的技术创新潜能, 本赛题不仅设立了丰厚的现金奖励, 更针对优秀的青年算法人才开辟了专属的职业发展绿色

通道。具体激励设置如下:

### (一) 赛事奖金设置

本赛题将根据总决赛的最终综合成绩(含线上客观评测与线下答辩),对全国总决赛一等奖前六名参赛团队按照如下标准颁发奖金:

1. 冠军奖:第 1 名,奖金 8000 元/每团队
2. 亚军奖:第 2-3 名,奖金 6500 元/每团队
3. 季军奖:第 4-6 名,奖金 3000 元/每团队

### (二) 人才招聘专属特权

为加速“人工智能+”创新成果向新质生产力的转化,切实打通产学研用链路,针对在本赛题中展现出卓越代码能力与工程落地思维的优秀参赛选手(不限于前六名获奖团队成员),出题方还将额外提供以下极具含金量的职业发展激励:

#### ● 核心岗位面试直通卡:

在复赛或总决赛中表现突出的核心算法选手,将获得人工智能研究院算法研发岗位的“校招面试直通卡”。凭此卡可直接免除在线笔试与简历初筛环节,直通终审技术面,优先锁定高潜算法人才核心席位。

#### ● 专项精英实习计划:

定向开放工业计算机视觉、大模型前沿应用等核心研发方向的为优质实习岗位。受邀实习生将直接参与真实的高级别产线视觉项目,享有充足的算力资源支持与资深算法研究员的 1v1 业务指导,积累宝贵的工业界实战经验,实习期间表现优异者可获全职 Offer 提前转正机会。

## 十五、其他说明

**公平性:** 严禁任何形式的作弊行为,包括但不限于数据泄露、模型预训练数据与测试数据重叠、抄袭他人代码等。一经发现,立即取消参赛资格,并追究相关责任。

**知识产权:** 参赛者提交的作品必须为原创,未在其他比赛中获奖或公开发表。比赛主办方有权对参赛作品进行展示、宣传等相关活动,但知识产权仍归参赛者所有。

## 十六、联系方式

赛题交流 QQ 群: 1092742972

邮箱: [aicomp2026\\_cut@163.com](mailto:aicomp2026_cut@163.com)

报名官网: [www.aicomp.cn](http://www.aicomp.cn)

# 低碳钢金相图像无监督相区分割

## 一、赛题背景

铁素体-珠光体钢（低碳钢）因强韧兼备的优异特性，占工业用钢总量 40%以上，广泛应用于汽车制造、机械装备、建筑工程等关键领域。晶粒度与晶粒形貌是决定该类钢材力学性能的核心微观参数，其大小、均匀性与边界完整性直接影响材料的强度、韧性及抗疲劳性，晶粒异常、混叠、边界模糊易导致核心零部件在服役过程中失效，引发安全风险。

在工业生产与检测场景中，晶粒度检测面临双重核心难题：一方面，传统检测依赖人工依据 GB/T6394-2017、ASTME112 等标准评级，不仅对检测人员的经验要求极高，且低碳组织中珠光体片层、铁素体晶界、多晶粒粘连交互复杂，主观误差达 $\pm 0.5$ 级以上，解译效率低下；另一方面，工业环境下金相样本制备成本高、周期长，实例级标注需专业人员逐晶粒勾勒，使得高质量标注数据获取难度大、成本高昂，难以满足大规模检测需求。学术与技术层面，无监督、小样本学习虽为突破数据瓶颈的关键方向，但现有算法在低碳钢多晶粒实例区分、晶界精准提取、粘连晶粒分割等方面仍存在技术短板，难以适配工业实际场景。

## 二、赛题应用场景

随着机器视觉与深度学习技术在材料微观表征领域的深度渗透，如何在工业数据获取难、标注成本高的现实约束下，基于无监督、小样本技术从珠光体+铁素体组织的金相图像中精准高效提取晶粒度信息，成为突破传统检测局限的关键。本次竞赛聚焦低碳钢晶粒度智能检测的核心需求，提供贴合工业实际的小样本金相图像数据集，旨在吸引材料科学、计算机视觉等领域的研究人员与技术爱好者参与，促进学术界与工业界的技术交流合作，推动晶粒度检测向低数据依赖、自动化、精准化升级，进而加速智能检测技术在钢铁生产、机械制造、汽车工业等关键领域的落地应用，为材料质量控制与工艺优化提供核心技术支撑。

## 三、赛题任务

**核心任务：**基于赛方提供的 500 张无标签低碳钢金相图像，开发无监督或半监督实例分割算法（半监督算法可自主标注部分训练数据，标注总量不超过无标签数据的 50%，即 $\leq 250$  张），实现铁素体晶粒、珠光体区域的像素级实例分割，区分每一个独立铁素体晶粒与每一个独立珠光体团簇，并输出实例索引掩码图+实例类别对应文件，用于计算实例分割 mIoU。

### 具体范围界定：

1. 执行类别感知的实例分割：

- 每个独立铁素体晶粒为一个实例；
- 每个独立珠光体团簇/区域为一个实例；
- 2. 输出必须包含：每个像素属于哪个实例+每个实例属于哪个类别；
- 3. 训练数据仅限官方 500 张无标签数据；
- 4. 输出格式严格匹配尺寸、坐标、命名规范。

#### 四、数据集及数据说明

##### （一）数据来源

数据来源于真实工业场景下的低碳钢试样，经标准磨抛、腐蚀工艺处理后，通过高倍显微镜（500 倍）拍摄获取，涵盖不同生产批次、腐蚀工艺的典型样本，具有广泛的工业代表性。

##### （二）相区定义及图像特征（标注参考）

###### 1. 铁素体晶粒（FerriteGrain）

类别：1

定义：碳在  $\alpha$ -Fe 中的间隙固溶体，低碳钢主要软相组织；单个铁素体晶粒为独立实例。

图像特征：在金相图像中呈亮白色/浅灰色，区域内部灰度均匀，无明显纹理结构，边界多为平滑曲线（晶界），单个晶粒尺寸通常在 5-20 $\mu\text{m}$  之间。

###### 2. 珠光体区域（PearliteRegion）

类别：0

定义：铁素体与渗碳体层片状共析组织，低碳钢主要硬相组织；连续珠光体团簇为独立实例。

图像特征：在金相图像中呈深灰色/黑色，具有明显的片层状纹理（放大后可见交替排列的明暗条纹），常以块状或带状分布，与铁素体区域形成显著灰度对比。

##### （三）精简标注说明

1. 标注工具：推荐使用 LabelMe、VGGImageAnnotator（VIA）等开源标注工具，支持像素级实例掩码标注与类别标记；

###### 2. 标注原则：

优先区分铁素体晶粒与珠光体区域核心类别，边缘模糊区域依据灰度/纹理判定（无片层纹理归为铁素体）；

每个独立铁素体晶粒单独标注为一个实例，每个连续珠光体团簇单独标注为一个实例；

粘连晶粒沿晶界趋势拆分，极小杂质点忽略并归为相邻实例；

###### 3. 标签赋值：

（1）例索引掩码：像素值为 1~255，每个实例赋予唯一编号，0 为背景/忽略；

(2) 实例类别：铁素体晶粒=1、珠光体=0，与实例编号一一对应；

(3) 输出格式：实例索引掩码图 (inst.png) +实例类别 JSON 文件，无多余数值干扰。

#### (四) 数据规模

无标签训练集包含 500 张图像，用于模型的无监督/半监督训练，数据中涵盖相区粘连、杂质干扰等真实工业场景；初赛测试集包含 68 张图像，用于模型性能的初步评估，赛方提供真实标签用于机器自动评分；复赛测试集包含 100 张图像，用于模型性能的进阶评估，同样提供真实标签用于评分；半决赛测试集包含 150 张图像，用于最终性能评测，其中包含弱腐蚀、晶粒异常等复杂工况样本；此外提供 1 张示例图像，用于展示数据格式与相区参考（标注铁素体与珠光体区域）。

#### (五) 数据格式及示例

1. 输入图像：JPG 格式，包含两种常见分辨率（1244×1044 像素、2448×2048 像素），3 通道（像素值范围 0-255），示例文件名：“train\_001.jpg”“test\_001.jpg”；

2. 分割标签/提交结果格式（PNG 掩码图）：

##### (1) 格式要求：

实例索引掩码图：8 位单通道 PNG 图像，分辨率与对应输入图像完全一致（需匹配输入图的 1244×1044 或 2448×2048 像素），像素值为 0~255，其中 0 表示背景/忽略，1~255 表示不同实例编号，每个独立铁素体晶粒、独立珠光体团簇均为唯一实例编号；

实例类别对应文件：JSON 格式，与图像同名，记录每个实例编号对应的类别，0=珠光体，1=铁素体晶粒。

##### (2) 命名规则：

实例索引掩码：输入图 test\_001.jpg→输出 test\_001\_inst.png

实例类别文件：输入图 test\_001.jpg→输出 test\_001\_class.json

##### (3) 示例说明（以局部区域为例）：

输入图像中坐标(100,200)至(150,250)的亮白色均匀铁素体晶粒→实例掩码对应区域像素值=1，JSON 中"1":1；

输入图像中坐标(160,210)至(200,260)的另一独立铁素体晶粒→实例掩码对应区域像素值=2，JSON 中"2":1；

输入图像中坐标(300,400)至(350,450)的深灰色片层珠光体团簇→实例掩码对应区域像素值=3，JSON 中"3":0。

验证方式：用 Python 的 PIL 库读取实例掩码后，需满足 mask.dtype=='uint8' 且 mask.shape==输入图.shape；实例编号范围为 0~255，JSON 文件中包含所有非 0 实例编号，且类别仅允许为 0 或 1，否则视为无效提交。

## （六）数据组织

训练集与测试集均按文件夹分类存储，文件夹结构示例：

数据集根目录/

```
├─train_images/#无标签训练集图像（JPG 格式）
│   ├──train_001.jpg
│   ├──train_002.jpg
│   └─...
├─test_images/#测试集图像（初赛/复赛/半决赛，JPG 格式）
│   ├──test_001.jpg
│   ├──test_002.jpg
│   └─...
└─example/#示例图像及标注参考
    ├──example_image.jpg
    └─example_results.png
```

## （七）数据特点

包含多种干扰场景：相区粘连、腐蚀不均、微小杂质、晶界模糊等；

铁素体与珠光体灰度差异存在波动（受腐蚀工艺影响），部分样本灰度重叠度较高；

无标签数据无类别偏倚，覆盖低碳钢典型组织比例（铁素体占比 6%-90%）；

图像分辨率包含 1244×1044、2448×2048 两种，需模型适配不同尺寸输入。

# 五、算法设计要求

## （一）模型类型

1. 核心要求：以无监督分割算法为基础，允许采用半监督学习（仅使用自主标注的训练数据，标注总量≤250 张）；

2. 禁止使用全监督预训练模型（如基于其他金相图像数据集预训练的模型），仅允许使用通用图像分割模型的无监督初始化（如 SAM 模型的无监督特征提取，禁止使用预训练的相区分割权重）；

3. 推荐算法方向：无监督聚类（如 K-Means 改进）、自监督特征学习（如对比学习）、生成式模型（如 VAE、扩散模型）、传统图像分割（如阈值分割、区域生长）的改进方案等。

## （二）创新性

1. 鼓励提出针对材料微观图像的无监督分割创新架构，如融合金相学先验知识（如相区灰度分布特性、晶界连续性）的算法；

2. 鼓励改进现有无监督算法的鲁棒性，适配工业场景中的灰度波动、杂质干扰

等问题；

3. 主动学习策略的创新性（如数据筛选逻辑、标注效率提升方法）将作为主观评分的重要依据。

### （三）可扩展性

1. 算法应具备良好的跨设备兼容性；
2. 模型需适配两种输入分辨率（1244×1044、2448×2048 像素），分割性能无显著下降；
3. 算法应具备低资源消耗特性，模型参数量不超过 500M。

### （四）数据使用限制

1. 仅允许使用赛方提供的 500 张无标签训练集及自主标注的训练数据，且自主标注数据总量不得超过无标签训练数据的 50%（即≤250 张）；
2. 严禁使用其他公开/私有金相图像数据集、外部标签或测试集标注信息；
3. 禁止对测试集数据进行人工标注或修改，严禁通过任何方式获取测试集真实标签；
4. 组委会将对代码及标注数据使用情况进行复现验证，若发现违规使用数据，取消参赛资格。

## 六、性能指标要求

本次竞赛采用交并比（IoU）和平均交并比（mIoU）作为核心评价指标，具体定义：

### 1. 实例分割 mIoU（占总分 50%）

#### （1）实例匹配规则

- a. 仅对类别相同的实例进行匹配；
- b. 一个真实实例最多匹配一个预测实例，一个预测实例最多匹配一个真实实例；
- c. 匹配依据：IoU 最大；

只有匹配  $\text{IoU} \geq 0.5$  的配对，才视为有效匹配。

#### （2）实例 IoU 计算

对一组有效匹配的真实实例与预测实例，实例 IoU 定义为：

$$(\text{IoU}_{inst} = \frac{\text{Area}(P \cap G)}{\text{Area}(P \cup G)})$$

P：预测实例的像素区域

G：真实实例的像素区域

∩：交集（同时属于预测与真实的像素数）

∪：并集（属于预测或真实的像素数）

IoU\_inst: 单个实例的分割精度

(3) 实例 mIoU 计算公式

$$(mIoU_{inst} = \frac{1}{N} \sum_{i=1}^N IoU_{inst}^{(i)})$$

N: 有效匹配的实例对总数

IoU\_inst(i): 第 i 个有效实例对的实例 IoU

mIoU\_inst: 所有有效实例的平均交并比, 为实例分割核心指标

(4) mIoU 得分 (满分 50 分)

$$(Score_{mIoU} = mIoU_{inst} \times 50)$$

2. 晶粒度面积法得分 (占总分 50%)

参考 GB/T6394-2017 平均晶粒度面积法, 做竞赛简化计算:

(1) 平均晶粒面积

真实铁素体晶粒平均面积:

$$(\bar{A}_{gt} = \frac{1}{M} \sum_{j=1}^M A_{gt}^{(j)})$$

预测铁素体晶粒平均面积:

$$(\bar{A}_{pred} = \frac{1}{K} \sum_{k=1}^K A_{pred}^{(k)})$$

M: 真实铁素体晶粒数量

K: 预测铁素体晶粒数量

A\_gt(j): 第 j 个真实铁素体晶粒的像素面积

A\_pred(k): 第 k 个预测铁素体晶粒的像素面积

(2) 面积相对误差

$$(E_{area} = \frac{|\bar{A}_{pred} - \bar{A}_{gt}|}{\bar{A}_{gt}})$$

(3) 面积法得分

$$(Score_{area} = \max(0, 1 - E_{area}) \times 50)$$

(满分 50 分, 误差越小得分越高)

3. 最终总分

$$(Total\ Score = Score_{mIoU} + Score_{area})$$

(满分 100 分, 两指标各占 50%)

## 七、功能要求

### (一) 准确性

对灰度波动、杂质干扰等复杂场景也能准确识别, 避免严重漏分割或误分割。

### (二) 可靠性

算法在不同批次、不同腐蚀工艺、不同分辨率的测试样本上性能稳定。

### (三) 鲁棒性

算法需抵御图像噪声、光照变化等干扰, 分割结果不受轻微图像预处理(如灰度归一化、平滑滤波)的影响;

禁止过度依赖特定图像特征(如固定灰度阈值), 需适配不同拍摄参数、不同分辨率下的样本。

### (四) 无监督兼容性

纯无监督方案需完全不依赖人工标注数据, 仅通过无标签数据实现相区分割;

半监督方案需明确标注数据的使用比例及作用(如仅用于微调、特征校准等), 标注数据量需符合 $\leq 250$ 张的限制。

## 八、开发环境

### (一) 编程语言

推荐使用 Python3.6 及以上版本, 需兼容主流数据处理与深度学习库。

### (二) 框架与工具

允许使用开源框架: PyTorch、TensorFlow、Scikit-learn、OpenCV 等;

允许使用通用无监督分割工具(如 SAM、CLIP 等)进行特征提取, 但禁止使用预训练的相区分割模型;

禁止使用商业闭源软件(如 MATLAB 的专有分割工具箱)或封装式 API 直接输出分割结果。

### (三) 硬件环境

开发阶段: 无强制要求, 可使用本地工作站或云端服务器;

测试阶段: 代码复现将在标准环境中运行, 确保公平性。

### (四) 依赖库

需提供完整的依赖库清单(requirements.txt), 明确版本号, 避免兼容性问题。

## 九、成绩评价

### (一) 评分构成

初赛和复赛采用机器自动评分(以实例分割综合得分为核心指标), 该成绩不计入决赛总分; 半决赛采用机器自动评分(以实例分割综合得分为核心指标), 成绩占总决赛总分的 70% (客观分); 总决赛采用主观评分(基于答辩表现), 成绩占总

决赛总分的 30%（主观分），包含算法创新性、工程性等评价维度。

## （二）评分细则

1. 客观评分（满分 70 分）：客观分为实例分割综合得分按比例折算，综合得分由实例 mIoU（50%）与晶粒度面积法得分（50%）加权组成。实例 mIoU 得分=实例 mIoU $\times$ 50；晶粒度面积法得分= $\max(0,1-\text{面积相对误差})\times 50$ ；实例分割综合得分=实例 mIoU 得分+晶粒度面积法得分（满分 100 分）；客观分=实例分割综合得分 $\times$ 0.7（满分 70 分）。

2. 主观评分（30 分）：算法创新性占 10 分（评价无监督实例分割、晶粒解缠、主动学习策略的创新程度）；工程落地性占 10 分（评价算法效率、分辨率适配能力、粘连/弱腐蚀样本鲁棒性）；技术报告质量占 5 分（评价方案阐述、实验分析、结果可视化）；答辩表现占 5 分（评价问题响应、技术细节解释）。

3. 加分项：纯无监督方案（未使用任何自主标注数据）且实例分割综合得分 $\geq$ 70 分，额外加 5 分（加分后总分不超过 100 分）。

## 十三、解题思路

### （一）知识点（含主动学习相关）

#### 1. 无监督分割核心知识点

（1）无监督聚类算法（K-Means、谱聚类）：考查基于灰度/特征的聚类优化能力，重点在于应对灰度波动场景的鲁棒性设计；

（2）自监督特征学习（对比学习、掩码自编码）：考查从无标签数据中提取判别性特征的能力，需结合金相图像的结构特性；

（3）传统图像分割（阈值分割、区域生长、形态学操作）：考查基础分割算法的改进与融合能力，重点在于边缘优化与噪声抑制。

#### 2. 主动学习相关知识：

（1）数据筛选策略（不确定性采样、核心集选择、多样性采样）：考查高价值数据的识别逻辑，需结合分割任务的误差分布特点；

（2）半监督学习融合（伪标签生成、一致性正则化）：考查少量标注数据的高效利用能力，重点在于减少标注偏差的影响；

（3）标注流程设计：考查标注效率与模型性能的平衡能力，需明确标注数据的迭代使用逻辑（标注总量 $\leq$ 250 张）。

#### 3. 工程实现知识点：

（1）图像预处理（灰度归一化、噪声过滤、增强）：考查适配工业图像质量差异的处理能力；

(2) 分辨率适配技术（如自适应缩放、多尺度特征提取）：考查模型对不同尺寸输入的兼容能力；

(3) 模型轻量化与效率优化：考查在有限硬件资源下的性能与速度平衡能力（针对 2448×2048 高分辨率图像）。

## (二) 思路引导（含主动学习方案）

### 1. 纯无监督思路：

先通过图像预处理（如直方图均衡化、高斯滤波）统一数据分布，减少腐蚀工艺差异带来的影响；

针对不同分辨率图像，可采用自适应缩放至统一尺寸处理（需注意保持相区比例），或设计多尺度特征提取模块适配原始尺寸；

基于自监督学习模型（如 MoCo、SimSiam）提取图像特征，再利用聚类算法（如改进 K-Means 加入空间约束）实现相区划分；

结合形态学操作（膨胀、腐蚀、边界平滑）优化分割结果，修正相区粘连与边缘模糊问题。

### 2. 主动学习+半监督思路：

初始化：用无监督算法（如传统分割+聚类）得到初始模型，对无标签数据（涵盖两种分辨率）进行预测；

数据筛选：基于预测不确定性（如熵值、置信度）或特征多样性，筛选高价值数据进行自主标注（累计≤500 张，建议两种分辨率数据均有覆盖）；

迭代优化：用标注数据训练半监督模型（如伪标签辅助、一致性正则化），再用新模型重新筛选高价值数据，迭代提升性能；

标注策略：可采用“小批量多次标注”（如每次 50 张），避免一次性标注过多低价值数据，确保标注效率。

### 3. 关键方向提示：

可融合金相学先验知识（如铁素体灰度范围、珠光体片层纹理特征）约束分割逻辑；

针对高分辨率图像（2448×2048），可采用分块处理或轻量化网络架构，平衡分割精度与速度；

针对复杂样本（弱腐蚀、晶界模糊），可设计多尺度特征融合或注意力机制提升鲁棒性。

## (三) 注意事项

### 1. 数据处理注意事项

避免过度预处理导致特征丢失（如强平滑可能模糊珠光体片层纹理）；

处理不同分辨率图像时，需保持相区位置与比例一致性，避免缩放导致的边缘错位；

标注数据需严格遵循“铁素体=1、珠光体=0”的规则，且总量不超过 250 张，避免违规使用标注数据。

### 2. 算法实现注意事项

纯无监督方案需避免隐性依赖标注信息（如手动设置固定阈值本质为人工标注）；

主动学习需记录完整的标注筛选逻辑与迭代过程，便于组委会复现验证；

模型需同时适配 1244×1044 和 2448×2048 分辨率，建议在训练中涵盖两种尺寸数据，提升泛化能力；

压缩包需按要求命名且无嵌套文件夹，否则可能导致评测系统无法读取；

提交代码需包含完整的训练与推理流程，明确分辨率适配方案，未提供可复现代码将影响成绩评定。

## 十一、赛题约束条件

### （一）算法约束

核心技术限定：以无监督分割算法为基础，允许采用半监督学习（需结合自主标注数据），但禁止使用基于其他金相图像数据集的全监督预训练模型；仅允许使用通用图像分割模型的无监督初始化（如 SAM 模型的无监督特征提取），禁止使用预训练的相区分割权重。

工具使用限制：允许使用开源框架（PyTorch、TensorFlow、Scikit-learn、OpenCV 等）及开源工具（如 LabelMe、VGGImageAnnotator），禁止使用商业闭源软件（如 MATLAB 的专有分割工具箱）或封装式 API 直接输出分割结果。

性能与效率约束：模型参数量不超过 500M；训练时间不超过 72 小时（普通工作站 CPU+单 GPU 环境）；需适配 1244×1044、2448×2048 两种分辨率输入，分割性能无显著下降。

### （二）数据使用约束

数据来源限定：仅允许使用赛方提供的 500 张无标签训练集图像及自主标注的训练数据，严禁使用其他公开/私有金相图像数据集、外部标签或测试集标注信息。

自主标注限制：半监督学习中自主标注的训练数据总量不得超过无标签训练数据的 50%（即≤250 张），需明确标注筛选逻辑与使用方式，便于组委会复现验证。

测试集使用规范：禁止对测试集数据进行人工标注、修改或通过任何方式获取真实标签；严禁将测试集数据用于训练过程。

数据保密要求：赛事提供的所有数据集（训练集、测试集、示例图像）仅可用于本次赛事相关研究与参赛，禁止泄露、传播或用于非赛事相关的商业用途，违反者取消参赛资格。

## 十二、参考资源

### （一）学术文献

《UnsupervisedImageSegmentationwithDeepLearning:ASurvey》

《ContrastiveLearningforUnsupervisedSegmentationofMedicalImages》

《主动学习在图像分割中的应用研究》

《金相图像分析中的无监督分割方法综述》

《多尺度特征融合在高分辨率图像分割中的应用》

### （二）开源工具

SAM（SegmentAnythingModel）：无监督特征提取；

Scikit-image：传统图像分割算法（阈值分割、区域生长）；

PyTorch-Spectral-Clustering：谱聚类实现；

DeepLab 系列：无监督/半监督分割框架；

LabelMe：开源像素级标注工具；

VGGImageAnnotator（VIA）：轻量级开源标注工具。

### （三）行业标准

GB/T6394-2017《金属平均晶粒度测定方法》

ASTME112《StandardTestMethodsforDeterminingAverageGrainSize》

## 十三、提交要求

### （一）初赛和复赛提交内容

1. 实例分割结果文件：测试集的实例索引掩码图（.png）+实例类别文件（.json），按以下规则命名：

实例掩码：test\_001\_inst.png

类别文件：test\_001\_class.json

压缩为：初赛(复赛)结果\_团队名.zip（压缩包内无嵌套文件夹）

2. 技术摘要：PDF 格式（≤2 页），简述算法思路、核心逻辑；

3. 提交限制：每天最多提交 5 次，排行榜实时更新，以最后一次提交为准。

### （二）半决赛、决赛提交内容

1. 分割结果文件：测试集的分割掩码图（PNG 格式），命名规则同初赛；

2. 技术报告：PDF 格式（≥1000 字），详述算法架构、创新点、实验结果、标注数据使用情况（含筛选逻辑、标注数量、使用方式）；

3. 模型文件：训练好的模型权重（格式不限，需支持复现）；

4. 完整代码：含数据预处理、模型训练、推理脚本，需提供详细注释及运行说明（README.md）；
5. 技术报告升级版：补充半决赛测试集实验分析，算法创新等等；
6. 答辩 PPT：≤10 页，提炼核心技术与成果。
7. 作品海报：尺寸为 A1（841mmx594mm），电子版分辨率为 300dpi，文件统一保存为 PNG 格式。

## 十四、奖金设置

为了鼓励参赛选手参赛积极性，激发在复杂制造场景下的技术创新潜能，本赛题不仅设立了丰厚的现金奖励，更针对优秀的青年算法人才开辟了专属的职业发展绿色通道。具体激励设置如下：

### （一）赛事奖金设置

本赛题将根据总决赛的最终综合成绩（含线上客观评测与线下答辩），对全国总决赛一等奖前六名参赛团队按照如下标准颁发奖金：

4. 冠军奖：第 1 名，奖金 8000 元/每团队
5. 亚军奖：第 2-3 名，奖金 6500 元/每团队
6. 季军奖：第 4-6 名，奖金 3000 元/每团队

### （二）人才招聘专属特权

为加速“人工智能+”创新成果向新质生产力的转化，切实打通产学研用链路，针对在本赛题中展现出卓越代码能力与工程落地思维的优秀参赛选手（不限于前六名获奖团队成员），出题方还将额外提供以下极具含金量的职业发展激励：

#### 1. 核心岗位面试直通车：

在复赛或总决赛中表现突出的核心算法选手，将获得人工智能研究院算法研发岗位的“校招面试直通车”。凭此卡可直接免除在线笔试与简历初筛环节，直通终审技术面，优先锁定高潜算法人才核心席位。

#### 2. 专项精英实习计划：

定向开放工业计算机视觉、大模型前沿应用等核心研发方向的优质实习岗位。受邀实习生将直接参与真实的高级别产线视觉项目，享有充足的算力资源支持与资深算法研究员的 1v1 业务指导，积累宝贵的工业界实战经验，实习期间表现优异者可获全职 Offer 提前转正机会。

## 十五、其他说明

### （一）公平性

严禁任何形式的作弊行为，包括但不限于使用测试集标签、抄袭他人代码、共享标注数据；

组委会将通过代码复现、IP 溯源等方式核查作弊行为，一经发现立即取消参赛

资格。

### （二）知识产权

参赛作品需为原创，未在其他比赛中获奖或公开发表；

参赛团队拥有作品的知识产权，组委会有权对作品进行展示、宣传；

优胜算法可与组委会协商进行产业化转化，相关权益另行约定。

### （三）更新与答疑

赛题相关更新（含最终确认的图像分辨率）将通过官网、QQ群发布，参赛团队需持续关注；

技术疑问可通过赛题交流QQ群提交，组委会将在2个工作日内回复。

## 十七、联系方式

赛题交流QQ群：1092664543

邮箱：[caitao@njsteel.com.cn](mailto:caitao@njsteel.com.cn)

报名官网：[www.aicomp.cn](http://www.aicomp.cn)

# 高炉铁次预测

## 一、赛题背景

高炉是钢铁生产中将铁矿石炼成铁水的核心设备。生产过程中，高炉会周期性出铁。一次从开口放铁到本次堵口结束的完整过程，称为一个“铁次”。本赛题中的“出铁量”是指该铁次最终产出的铁水总量，单位为吨；“出铁时长”是指该铁次从开铁到堵口完成所经历的总时间，单位为分钟。

在实际生产中，调度人员需要在铁次开始时，提前判断本次大约能出多少铁、会持续多久，以便安排铁水罐、运输和炼钢接铁计划。如果判断偏差较大，就可能出现准备不足、等待时间增加或资源占用不合理等问题。

因此，本赛题要解决的问题可以简单理解为：给定某钢铁企业高炉在目标铁次开始前已经能够看到的历史数据，建立模型预测该铁次最终会产出多少铁水，以及这次出铁会持续多长时间。赛题重点考察数据理解、时间对齐、特征构造和回归预测能力。

## 二、赛题应用场景

本赛题的应用场景是钢铁企业的高炉出铁调度。每当一个新铁次即将开始时，现场都需要提前准备好接铁和运输资源。如果能够提前判断该铁次的出铁量和持续时间，就能更合理地安排铁水罐、运输和炼钢接收计划；如果判断不准，就可能出现等待、拥堵或准备不足。

本赛题聚焦单铁次预测场景。参赛者需要在目标铁次开铁时，基于此前已经积累的历史数据，预测该铁次结束后的两个结果，即总出铁量和总出铁时长。该任务可以作为后续生产组织、运输调度和计划优化的基础模块。

## 三、赛题任务

参赛者需要利用主办方提供的高炉历史生产数据，建立单铁次预测模型。在每个待预测样本中，只能使用该样本预测参考时刻及之前已经产生的数据，预测目标铁次最终的出铁量和出铁时长。

### （一）任务描述

本赛题的核心任务是：针对样本主表中给出的每一个目标铁次，基于官方提供的样本索引表、历史出铁实绩、小时级炉况过程数据和事件级变料/质量数据，自主完成数据清洗、时序对齐、特征构造、模型训练和结果推理。训练样本时间范围为2024年3月1日至2024年11月30日；初赛测试样本时间范围为2024年12月1日至2024年12月31日；复赛测试样本时间范围为2025年1月1日至2025年1月31日；半决赛测试样本时间范围为2025年2月1日至2025年3月26日。

### （二）任务输入输出说明

1. 输入数据：训练/测试样本主表（`train_samples.csv`、`test_a_samples.csv`、

test\_b\_samples.csv、test\_c\_samples.csv），以及与之配套的小时级炉况过程表（operation\_hourly.csv）、事件级变料/质量表（burden\_change.csv）和训练期历史出铁实绩表（tap\_history\_train.csv）。

2. 输入时间与约束：官方提供的全部样本时间范围覆盖 2024 年 3 月 1 日至 2025 年 3 月 26 日。样本主表中的预测参考时刻统一为目标铁次的开铁时间。对任一训练样本或测试样本进行预测时，仅允许使用该参考时刻及其之前的数据，禁止使用本铁次后续过程信息和任何未来信息。

3. 输出结果：对每个目标铁次输出两个连续值，分别为该铁次结束后的总出铁量预测值（单位：吨）和该铁次从开铁到堵口完成的总时长预测值（单位：分钟）。

4. 任务属性：这是一个面向真实工业场景的多表时序回归任务，既需要完成多源数据的时间对齐，也需要对两个连续目标进行联合预测或分别预测。

#### 四、数据集及数据说明

##### （一）数据来源

数据来源于某钢铁企业某大型高炉连续生产过程中的真实业务记录，主要包括三类数据：出铁实绩数据、小时级炉况过程数据和事件级变料/质量数据。原始数据经主办方进行匿名化、脱敏、字段标准化和格式转换后，以赛事专用数据包形式提供。

##### （二）数据规模

本赛题所有样本时间范围覆盖 2024 年 3 月 1 日至 2025 年 3 月 26 日。原始小时级炉况数据约 9360 条，事件级变料记录约 4021 条，出铁实绩明细约 22004 条。由于同一铁次可能对应多条铁罐明细记录，官方将其按铁次主键进行聚合，最终形成约 3959 个铁次级样本。

本赛题采用时间顺序切分方式组织数据，以贴近工业场景中的实际部署方式并避免时间穿越。训练集时间范围为 2024 年 3 月 1 日至 2024 年 11 月 30 日，共约 2754 个标注样本；初赛测试集时间范围为 2024 年 12 月 1 日至 2024 年 12 月 31 日，共约 335 个样本；复赛测试集时间范围为 2025 年 1 月 1 日至 2025 年 1 月 31 日，共约 322 个样本；半决赛测试集时间范围为 2025 年 2 月 1 日至 2025 年 3 月 26 日，共约 548 个样本。官方仅提供训练集标签，测试集不提供标签，参赛者可在训练集内部自行划分验证集。

##### （三）数据文件说明

1. train\_samples.csv（铁次级）：训练样本主表，包含样本编号、铁次号、铁口号、预测参考时刻、出铁量标签、出铁时长标签等字段；

2. test\_a\_samples.csv（铁次级）：初赛测试样本主表，包含样本编号、铁次号、铁口号、预测参考时刻等字段，不含标签；

3. test\_b\_samples.csv（铁次级）：复赛测试样本主表，字段与初赛测试样本一致，

不含标签；

4. `test_c_samples.csv`（铁次级）：半决赛测试样本主表，字段与初赛测试样本和复赛测试样本一致，不含标签；

5. `operation_hourly.csv`（小时级）：高炉小时级过程数据，包含入炉风量、冷风压力、热风压力、富氧、热风温度、喷煤量、湿度、煤气利用率、炉顶压力、差压、风压比、炉顶温度平均、实际风速、炉喉温度等字段；

6. `burden_change.csv`（事件级）：高炉变料与质量数据，包含铁量、综合品位、矿耗、燃料比、焦比等事件级特征字段；

7. `tap_history_train.csv`（铁次级）：由出铁实绩明细聚合得到的训练期历史铁次实绩表，包含参考时刻、堵口时间、实际出铁量、出铁时长、受铁罐数等字段；

8. `data_dictionary.xlsx`（说明文件）：字段释义、单位说明、空值含义、时间字段口径及使用注意事项。

除 `data_dictionary.xlsx` 外，其余赛事数据均采用 CSV 格式发布，统一使用 UTF-8 编码；时间字段统一采用北京时间（UTC+8），数值字段保留原始工程单位。训练样本主表、测试样本主表和历史铁次实绩表中的参考时刻均对应目标铁次的开铁时间。参赛者需要利用该时刻及之前的数据，预测该铁次结束后的总出铁量和总出铁时长。对于少量缺失值，赛事方不进行统一插补，由参赛者自行设计处理方案。

#### （四）数据预处理说明

1. 对涉企字段、设备编码、罐号等敏感信息进行匿名化和脱敏处理；
2. 对异常时间戳、重复表头行等明显格式问题进行基础清洗；
3. 对出铁实绩明细按铁次主键进行聚合，保留总出铁量和总出铁时长作为训练标签；

4. 官方仅完成基础格式处理，不提供最终建模特征，参赛者需自行完成多表时序对齐和特征构建。

## 五、算法设计要求

### （一）模型类型

本赛题对算法类型不做严格限定，鼓励参赛者采用适用于工业结构化时序数据的机器学习或深度学习方法，包括但不限于 XGBoost、LightGBM、CatBoost、随机森林、TabNet、多层感知机、Transformer、多任务学习模型及集成学习方案。

### （二）创新性

鼓励参赛者围绕多表时序对齐、工业机理先验注入、异常样本鲁棒处理、多目标协同建模、结果校准和可解释性分析等方向形成创新方案。

### （三）可扩展性

算法方案应具备良好的可扩展性，能够在不同配置的 CPU/GPU 环境中运行，并

在样本规模增加、字段扩展或工况变化时保持相对稳定的性能表现。

#### （四）工程可落地性

算法方案应兼顾预测精度与工程可落地性，避免依赖不可复现的复杂人工规则或不可离线运行的外部服务。进入总决赛的方案需支持在统一环境下完整复现。

### 六、性能指标要求

#### （一）主要指标

1. 出铁量加权绝对百分比误差： $WMAPE\_iron = \sum | \text{预测出铁量} - \text{实际出铁量} | / \sum \text{实际出铁量}$ 。

2. 出铁时长加权绝对百分比误差： $WMAPE\_time = \sum | \text{预测出铁时长} - \text{实际出铁时长} | / \sum \text{实际出铁时长}$ 。

3. 综合机器评分： $Score\_main = 100 \times [1 - 0.5 \times WMAPE\_iron - 0.5 \times WMAPE\_time]$ ，最低记为 0 分，分值越高越优。

#### （二）次要指标

1. 出铁量平均绝对误差，用于辅助评估吨位偏差。
2. 出铁时长平均绝对误差，用于辅助评估节拍偏差。
3. 模型文件大小和全量测试推理时长，用于总决赛工程评价。

#### （三）目标参考

高质量方案应在隐藏测试集上稳定优于官方基准线方案，并在出铁量误差控制、出铁时长误差控制和综合机器评分等方面达到较优水平。

### 七、功能要求

#### （一）准确性

算法应能够为每个测试样本输出完整、数值合理的出铁量和出铁时长预测结果，并尽量降低大偏差样本的比例。

#### （二）可靠性

面对不同月份、不同料制、不同风氧制度和不同生产节奏区间的样本，算法应保持稳定输出，不应因局部工况变化而出现成批失真。

#### （三）鲁棒性

算法应对缺失值、重复值、异常值、统计口径差异和明细级多行聚合问题具备较强鲁棒性。

#### （四）时效性

在统一测评环境中，算法应在规定时间内完成全量测试样本推理。原则上，单样本平均推理耗时不超过 1 秒，全量阶段测试集推理时长不超过 30 分钟；正式阈值以赛事发布稿和测评细则为准。

### （五）可解释性

鼓励参赛者提供特征重要性分析、样本级误差分析或关键影响因素可视化结果，以增强方案的可解释性。

### （六）功能测试方式

赛事方将在统一离线环境中测试算法是否能够正确读取数据、按要求生成结果文件，并检查结果完整性、字段合法性、样本覆盖率、运行时长和可复现性。

## 八、开发环境

### （一）软件环境

1. 操作系统：Linux（推荐 Ubuntu20.04/22.04）或 Windows10/11。
2. 编程语言：Python3.10 及以上版本。
3. 推荐依赖：NumPy、Pandas、SciPy、scikit-learn、XGBoost、LightGBM、CatBoost、PyTorch、Matplotlib、Seaborn。
4. 工程文件：应提供 requirements.txt、README.md、运行脚本等辅助文件。

### （二）硬件环境

1. CPU 环境：8 核及以上 CPU，32GB 及以上内存。
2. GPU 环境：非强制要求；若采用深度学习方案，建议单卡显存 16GB 及以上。
3. 存储空间：建议可用磁盘空间 50GB 及以上。

## 九、成绩评价

### （一）指标权重

机器评分阶段，出铁量指标权重为 50%，出铁时长指标权重为 50%。排行榜及阶段晋级统一以综合机器评分为主排序指标。

### （二）评分细则

1. 初赛阶段：依据初赛测试集综合机器评分进行排序，用于开发验证和复赛晋级选拔。
2. 复赛阶段：依据复赛测试集综合机器评分进行排序，用于半决赛晋级选拔。
3. 半决赛阶段：依据半决赛测试集综合机器评分进行排序，并作为总决赛客观评分基础。
4. 若综合机器评分相同，则依次比较出铁量加权绝对百分比误差、出铁时长加权绝对百分比误差；若仍相同，则以较早提交时间为先。
5. 有效成绩阈值设定为：综合机器评分不低于 60 分，且不低于官方基线参考分数。

### （三）其他评分因素

总决赛采用“70%客观评分+30%主观评分”的评价方式。其中客观评分为半决赛

机器测评得分标准化结果，主观评分由专家组结合技术创新性、工业可落地性、结果复现性、技术文档质量和答辩表现综合给出。

1. 客观评分（机器测评）（权重：70%）：以半决赛隐藏测试集综合机器评分标准化结果计入。

2. 技术创新性（权重：10%）：主要考察特征设计、模型方法和问题解决思路。

3. 工程复现性（权重：10%）：主要考察代码运行、结果复现、效率和提交规范性。

4. 技术文档与答辩质量（权重：10%）：主要考察文档完整性、表达逻辑、分析深度和答辩表现。

## 十、解题思路

### （一）知识点

本赛题主要考核工业多表数据理解、时间序列对齐、特征工程、回归建模、多目标优化、异常处理、时间切分验证、结果解释与工程复现等能力。

### （二）思路引导

1. 明确铁次级标签口径，处理好出铁实绩明细中“同一铁次多行铁罐记录”的聚合问题。

2. 围绕样本参考时刻完成小时级炉况数据和事件级变料数据的时序对齐。

3. 重点关注入炉风量、风压、富氧、热风温度、喷煤量、湿度、煤气利用率、差压、风压比、炉顶温度、实际风速、炉喉温度、铁量、综合品位、矿耗、燃料比、焦比等关键变量。

4. 建议采用时间顺序切分进行验证，避免随机切分造成未来信息泄露。

5. 建议对大偏差样本进行分段分析，识别异常工况、极端铁量样本和长时长样本。

### （三）注意事项

1. 严禁在特征构建或验证中使用参考时刻之后的数据。

2. 需特别注意变料数据的事件级特性和时间窗口滞后效应。

3. 若采用深度学习方案，应注意不同字段量纲差异带来的归一化问题。

4. 结果文件中样本编号不得缺失、重复或错位。

## 十一、赛题约束条件

### （一）算法约束

1. 禁止使用任何商业闭源在线推理接口或需要联网调用的模型服务。

2. 禁止通过人工逐条修正测试结果。

3. 禁止使用未来信息、测试标签泄露信息或通过测试集反推标签。

4. 进入总决赛的方案须支持离线环境完整运行。

## （二）数据使用约束

1. 参赛者仅可使用赛事官方提供的数据，严禁使用任何外部公开数据、私有数据或自行采集数据。
2. 严禁使用依赖外部数据训练得到的预训练模型权重；允许使用开源算法框架、开源代码库和通用基础依赖库，但不得借助外部数据增强模型能力。
3. 禁止传播、转售、公开展示、用于商业用途或泄露赛事数据。
4. 所有数据仅授权用于本赛事相关算法研究与评测。

## 十二、参考资源

### （一）文献资料

1. XGBoost: A Scalable Tree Boosting System. KDD2016.
2. LightGBM: A Highly Efficient Gradient Boosting Decision Tree. NeurIPS2017.
3. CatBoost: unbiased boosting with categorical features. NeurIPS2018.
4. 工业过程数据建模、时间序列回归、多任务学习与异常检测相关公开论文和技术报告。

### （二）在线资源

1. 吴恩达《Machine Learning Specialization》（Coursera 视频课程）。
2. Google《Machine Learning Crash Course》（含讲解视频与练习）。
3. scikit-learn 官方 MOOC《Machine learning in Python with scikit-learn》。
4. 李宏毅《Machine Learning 2023 Spring》课程页面（含课程视频与课件）。
5. fast.ai《Practical Deep Learning for Coders》（公开视频课程）。

## 十三、提交要求

### （一）初赛、复赛及半决赛提交内容与格式

参赛者需将预测结果保存为 CSV 文件，并压缩为 ZIP 文件后提交。结果文件字段要求如下：

1. `sample_id`（字符串/整数）：与官方测试样本中的样本编号完全一致。
2. `pred_tap_iron`（浮点数）：预测出铁量，单位为吨。
3. `pred_tap_time_len`（浮点数）：预测出铁时长，单位为分钟。

CSV 文件应使用 UTF-8 编码，字段名称须与赛事要求完全一致；每个样本编号只能出现一次；不得缺行、重复或附加多余字段。预测值建议保留三位及以上小数。

注：提交限制：每天最多提交 5 次，排行榜实时更新，以最后一次提交为准。

### （二）半决赛、总决赛提交内容

1. 算法代码：包含数据读取、特征构建、模型训练、模型加载和预测推理等完整代码。
2. 模型文件：训练好的模型参数文件或可复现训练脚本。

3. 技术报告：PDF 或 Word 格式，不少于 3000 字。
4. README 文件：说明运行环境、依赖安装、运行命令和输入输出目录结构。
5. requirements.txt 或 environment.yml 文件。
6. 如赛事组织方需要，可额外提交演示视频、答辩 PPT 和补充说明材料。

### （三）提交规范

1. 所有提交材料须保证可读取、可解压、可运行。
2. 初赛、复赛和半决赛提交的 ZIP 压缩包内须包含且仅包含一个结果文件，结果文件固定命名为 result.csv。
3. 初赛压缩包命名规则为 teamname\_bf\_tap\_predict\_prelim.zip，复赛压缩包命名规则为 teamname\_bf\_tap\_predict\_round2.zip，半决赛压缩包命名规则为 teamname\_bf\_tap\_predict\_semifinal.zip。
4. 若代码与结果无法复现，赛事方有权取消成绩或奖项。
5. 最终解释权归赛事主办方所有。

## 十五、奖金设置

为鼓励参赛者积极参与并支持优秀成果落地，本赛题设置赛事奖金，并为表现突出的选手提供相应的职业发展支持。具体安排如下：

### （一）赛事奖金设置

本赛题将根据总决赛的最终综合成绩（含线上客观评测与线下答辩），对全国总决赛一等奖前六名参赛团队按照如下标准颁发奖金：

1. 冠军奖：第 1 名，奖金 8000 元/每团队
2. 亚军奖：第 2-3 名，奖金 6500 元/每团队
3. 季军奖：第 4-6 名，奖金 3000 元/每团队

### （二）优秀选手职业发展支持

针对在本赛题中表现突出的优秀选手，出题方还将提供相应的职业发展支持，具体包括以下内容：

#### 1. 核心岗位面试直通机会：

在复赛或总决赛中表现突出的核心算法选手，可获得人工智能研究院算法研发岗位的校招优先面试机会。

#### 2. 专项精英实习计划：

定向开放工业计算机视觉、大模型前沿应用等核心研发方向的优质实习岗位。受邀实习生可直接参与真实工业项目，并获得相应的业务指导与实践机会。

## 十六、其他说明

### （一）更新与答疑

赛题内容、数据说明、评测细则和基线方案如有更新，以赛事官网和官方答疑群

发布的信息为准。

### （二）公平性

严禁任何形式的作弊行为，包括但不限于数据泄露、标签泄露、抄袭他人代码、伪造结果和利用非授权数据获取不公平优势等。

### （三）知识产权

参赛者提交的作品应保证原创，不得侵犯任何第三方知识产权。比赛主办方有权在赛事宣传、成果展示和学术交流中使用参赛作品的相关内容，具体知识产权归属以赛事公告和双方约定为准。

### （四）数据保密与合规

赛事数据为脱敏后的工业数据，仅授权用于本赛事算法研究与评测。参赛者应承诺履行保密义务，不得传播、复用或用于与赛事无关的商业活动。

### （五）结果可复现性

进入总决赛的团队需保证提交结果可复现。若官方在指定环境中无法按照提交说明复现主要结果，赛事方有权取消其成绩或奖项。

## 十七、联系方式

赛题交流 QQ 群：1092645019

邮箱：liupeng.nanjing@gmail.com

报名官网：www.aicomp.cn

## 煤气发电量预测与发电优化

### 一、赛题背景

煤气是钢铁企业高炉、焦炉、转炉生产过程中的主要副产能源。这些煤气经回收净化后，可直接送入燃气锅炉和发电机组进行自发电，即企业利用副产煤气自己发电，从而减少从电网购电。与余热、光伏等其他自发电形式不同，煤气自发电的燃料来源于生产节奏不稳定的冶炼过程，其发生量波动大、可存储性差，只能依靠气柜进行短期缓冲，因此预测难度最高，也最直接地影响企业的整体用电成本。

煤气自发电的预测精度对企业购电策略至关重要。大多数钢铁企业参与电力现货市场交易，企业需要在今天根据次日预测的自发电量以及生产过程计划用电量，决定从电网购买多少电量。如果买少了，高峰时段就必须以高价临时补电；如果买多了，低价时段购入的多余电力既无法存储，也无法按购入价回售，造成成本浪费。更关键的是，企业向电网申报的最大需量直接决定了每月的基本电费支出。如果自发电量预测值偏小，实际自发电不足会导致企业从电网紧急取电，最大需量会飙升，次月电费剧增；如果预测值偏大，企业虚占需量容量同样造成浪费。煤气自发电量预测的精度直接决定了企业购电计划的合理性与经济性，预测误差直接影响月度成本波动。因此，高精度的煤气自发电量预测是企业降本增效的核心前提之一。

然而，仅做发电量预测远远不够。在实际生产中，发电调度并非“能发多少就发多少”，而是必须在多种约束下制定发电计划。第一是气柜容量约束：煤气柜有安全上下限，为保证储气用气安全，柜位过高会引发放散，既浪费能源又面临环保处罚；柜位过低则可能无法满足发电机组瞬时需求。第二是机组出力特性：燃气机组有最小稳定负荷和爬坡速率限制，不可随意启停。第三是电价峰谷约束：江苏省将一天划分为尖、峰、平、谷时段，企业希望在电价高时多发自发电来替代外购电，在电价低时适当少发，利用低价电网电。

本赛题旨在挖掘面向能源降本增效的智能算法方案。参赛者需要利用历史数据建立高精度的煤气自发电量预测模型；再在满足生产约束的前提下，设计合理的发电计划优化策略，明确机组何时启停、负荷如何分配，最终实现企业能源运营收益最大化。

### 二、赛题应用场景

本赛题的应用场景为钢铁企业能源调度中心。每日，调度员需完成以下工作：

1. 根据次日生产计划（铁水产量、轧钢节奏等）和历史数据，预测次日各时段的煤气自发电量。
2. 在气柜安全限值、机组爬坡约束、电价峰谷时段等条件下，制定未来 24 小时的发电机组启停计划和负荷分配方案。
3. 将发电计划与企业购电策略联动，向电网申报次日最大需量和购电曲线。

目前多数企业依赖人工经验进行调度，预测误差大、计划保守。本赛题旨在通过数据驱动的智能算法，实现预测与优化一体化，显著提升能源经济性。

### 三、赛题任务

#### （一）任务描述

参赛者需要掌握时间序列分析方法，能够进行多源异构数据的预处理与特征工程，熟悉线性规划、整数规划或启发式算法，利用提供的多源时序数据（煤气产耗、气柜状态、机组负荷、电价时段），完成以下核心任务：

1. 自发电量预测：构建高精度预测模型，输出未来一段时间的煤气产耗量及由此产生的发电机组负荷（即自发电量），评估指标为 1-MAPE。该预测对象是指，按照企业历史运行方式，煤气产耗趋势下的发电量。短周期指 2 小时，长周期指 24 小时。

2. 发电计划优化：基于预测结果获得的未来资源边界，在满足气柜容量、煤气用户需求等约束下，设计合理的发电机组调度策略（何时启停、负荷分配），以最大化峰谷电价收益。

3. 两项任务顺序执行，先完成预测，获得未来资源约束（例如可用于发电的煤气量），再基于这些约束进行优化。

#### （二）任务输入输出说明

1. 输入：历史数据文件，以及江苏省电价峰谷时段划分和电价数据。预测时仅允许使用参考时刻及其之前的数据，禁止未来信息泄露。

2. 输出：对每个测试样本（时间点或时段），输出两个预测值：单台机组负荷（若有多台）和总发电负荷（具体以官方提交模板为准），以及输出一个未来一段时间的发电计划（各煤气量）。

### 四、数据集及数据说明

#### （一）数据来源

数据来源于某钢铁企业高炉、焦炉、转炉生产过程中的真实业务记录，包括以下四类数据：

数据类型	描述	采样频率
高炉、焦炉、转炉的煤气发生量与消耗量	各炉煤气实时产耗数据	1 分钟~15 分钟
煤气柜数据	高炉煤气、焦炉煤气、转炉煤气的气柜流量、容量和压力	1 分钟~15 分钟
发电机组负荷	发电机组实时负荷数据	1 分钟~15 分钟
江苏省电价峰谷时段划分和电价	电价时段划分说明	-

原始基础数据经主办方进行匿名化、脱敏化、字段标准化和格式转换后，以赛事专用数据包形式提供。

## （二）数据规模

本赛题所有样本时间范围覆盖 2025 年 1 月 1 日至 2025 年 10 月 31 日，均为分钟级生产数据。

本赛题采用时间顺序切分方式组织数据，以贴近工业场景实际部署方式并避免时间穿越。赛事按阶段逐步释放数据，各阶段所释放的数据均可自由组合用于模型迭代与训练。数据集划分如下：

初赛阶段：2025 年 1 月 1 日至 2025 年 4 月 30 日（约 4 个月，涵盖冬春交替工况）。评测数据（指用于评分、选拔等）：2025 年 5 月 1 日至 2025 年 5 月 2 日。

复赛阶段：2025 年 5 月 3 日至 2025 年 7 月 31 日数据（约 3 个月，涵盖夏季高负荷工况）。评测数据：2025 年 8 月 1 日至 2025 年 8 月 5 日。

半决赛阶段：2025 年 8 月 6 日至 2025 年 9 月 30 日数据（约 2 个月，涵盖夏秋交替工况）。评测数据：2025 年 10 月 1 日至 2025 年 10 月 10 日。

决赛阶段：测评数据 2025 年 10 月 11 日至 2025 年 10 月 31 日数据。

时间序列数据文件格式为 CSV（逗号分隔值），编码为 UTF-8，所有时间序列数据的时间戳格式统一为 YYYY-MM-DD HH:MM:SS。

数据集包含以下文件：

文件名称	数据粒度	主要说明
gas.csv	分钟级	包含高炉、焦炉、转炉煤气的发生量和消耗量等字段。
gas_holder.csv	分钟级	包含高炉、焦炉、转炉煤气气柜容量等字段。
gas_user.csv	分钟级	包含高炉、焦炉、转炉煤气用户的消耗量等字段。
load.csv	分钟级	包含发电负荷等字段。
江苏省电价峰谷时段	-	江苏省 1 月至 12 月电价时段划参考数据
data_dictionary.xlsx	说明文件	字段释义、单位说明及使用事项

## （三）数据字段说明

企业拥有多座高炉、焦炉和转炉，且煤气用户较多，为了便于参赛者快速理解数据，将以数字编号命名，数字编号不区分先后和用户优先级，例如“blast\_furnace\_1”表示 1 号高炉，“blast\_furnace\_user1”表示高炉煤气用户 1。

文件名称	参数	数据说明
gas.csv	blast_furnace_1	1 号高炉煤气发生量。
	air_heater_1	1 号热风炉的高炉煤气消耗量。热风炉作为高炉冶炼的重要附属设备，为高炉冶炼提供热量，但与此同时也消耗了高炉煤气。热风炉编号与高炉编号一一对应，例如 1 号热风炉只会消耗 1 号高炉生产的煤气。
	coke_oven_1	1 号焦炉煤气发生量。

	converter_1	1号转炉煤气发生量。
	into_gas_mixed_blast_furnace	进混气站的高炉煤气量。
	into_gas_mixed_coke	进混气站的焦炉煤气量。
	into_gas_mixed_converter	进混气站的转炉煤气量。
gas_holder.csv	blast_furnace_gas_holder_1	1号高炉气柜。气柜的作用是存储高炉煤气和稳定管网压力，可存储多个高炉的煤气。
gas_user.csv	blast_furnace_user1	高炉煤气用户1。消耗高炉煤气，需要优先供应高炉煤气，以保证正常生产。
	converter_user1	转炉煤气用户1。消耗转炉煤气，掺烧转气的目的是保证生产排放达标。
	mixed_gas_user1	混气用户1。除了直接使用煤气的用户外，还有些用户需要使用混合煤气（高炉、焦炉、转炉煤气混合），需要优先供应混合煤气，以保证正常生产。
load.csv	generator_1	发电机组1。
	generator_all	全部发电机组。6套机组总负荷，包含4套50MW和2套120MW发电机组。
	generator_use_coke_gas	发电机组消耗的焦炉煤气量。
	generator_use_converter_gas	发电机组消耗的转炉煤气量。
	generator_use_blast_furnace_gas	发电机组消耗的高炉煤气量。
price.xlsx	-	一年每月分时电价，列名为1月至12月。电价会随市场变化，该数据仅供本赛题使用。

**(四) 数据预处理说明**

原始数据已经过脱敏和基础格式统一（时间戳格式 YYYY-MM-DD HH:MM:SS，编码 UTF-8），但仍可能存在缺失值、重复值、异常值。参赛者需自行完成多表时序对齐、缺失处理、异常检测等清洗工作。

**五、算法设计要求**

**(一) 模型类型**

本赛题禁止使用任何闭源和商用大模型（包括但不限于 GPT-4、Claude、Gemini、文心一言商业版等），鼓励参赛者采用适合工业时序数据的机器学习或深度学习方法，包括但不限于 XGBoost、LightGBM、随机森林、多层感知机、Transformer 等。

## （二）创新性

本鼓励参赛者围绕多表时序对齐、工业机理先验注入、异常样本鲁棒性处理、多目标协同建模、可解释性分析等方向形成创新方案。

## 六、性能指标要求

本赛题聚焦煤气发电量预测。针对高精度和高响应的业务需求，本赛题采用 1-平均绝对百分比误差（MAPE）作为评估参赛模型性能的核心指标，其中 MAPE 的计算公式为：

$$MAPE = \frac{100\%}{n} \sum_{i=1}^n \left| \frac{y_i - \hat{y}_i}{y_i} \right| \in [0, +\infty)$$

参赛者的算法模型必须在保证基础准确率的前提下，尽可能保证响应速度。

本赛题同时也聚焦煤气发电优化。针对发电计划优化，本赛题采用约束满足性、电价时段利用效率和经济效益提升三个指标作为评估参赛模型的核心指标。其中，各指标的定义如下：

### 1. 约束满足性

本场景的气柜容量为 20 万  $\text{m}^3$ ，为了保证生产安全，气柜的柜位（柜容）全程保持在安全区间内，低柜位（下限）约为 15% 柜容，高柜位（上限）约为 90% 柜容，高高柜位约为 95% 柜容；煤气用户无供气不足的时段。发电机组有装机容量，正常运行下发电负荷在 60% 至 100% 间，场景中的发电机组为 4 套 50MW 和 2 套 120MW 发电机组，其中 generator\_1 指的是 4 套 50MW 机组的负荷，generator\_all 指的是 6 套机组总负荷。本赛题中，将机组的爬坡速率设定为 10% 额定容量/分钟。

### 2. 电价时段利用效率

电价加权负荷比的计算公式为：

$$E = \frac{\sum_t L_t \cdot p_t}{\sum_t L_t \cdot \bar{p}}$$

其中  $L_t$  为  $t$  时刻计划发电负荷， $p_t$  为  $t$  时刻电价（尖峰平谷对应不同单价）， $\bar{p}$  为全天平均电价。

### 3. 经济效益提升

将发电计划与历史发电进行经济效益对比，收益计算包括替代外购电效益（自发电每度替代外购电一度）和煤气放散惩罚，计算相对收益提升率：

$$\text{相对收益提升率} = (\text{优化效益} - \text{基准效益}) / \text{基准效益}$$

## 七、功能要求

### 1. 可靠性

面对不同月份和不同节奏区间的样本，算法应保持稳定输出，不应因局部工况变化出现成批失真。

### 2. 鲁棒性

算法应对缺失值、重复值、异常值、统计口径差异和明细级多行聚合问题具备较强鲁棒性。

### 3. 时效性

在统一测评环境中，算法应在规定时间内完成全量测试样本推理。原则上，单样本平均推理耗时不超过 30 秒，全量阶段测试集推理时长不超过 30 分钟；正式阈值以赛事发布稿和测评细则为准。

### 4. 可解释性

鼓励参赛者提供特征重要性分析、样本级误差分析或关键影响因素可视化，以增强结果可解释性。

### 5. 功能测试方式

赛事方将在统一离线环境中测试算法是否能够正确读取数据、按要求生成结果文件，并检查结果完整性、字段合法性、样本覆盖率、运行时长和可复现性。

### 6. 优化调度

输出未来一段时间的发电计划。

## 八、开发环境

### （一）软件环境

操作系统：Linux（推荐 Ubuntu 20.04/22.04）或 Windows 10/11。

编程语言：Python 3.8 - Python 3.10。

推荐依赖：NumPy、Pandas、SciPy、scikit-learn、XGBoost、LightGBM、CatBoost、PyTorch、Matplotlib、Seaborn。

工程文件：应提供 requirements.txt、README.md、运行脚本等辅助文件。

### （二）硬件环境

CPU 环境：8 核及以上 CPU，32GB 及以上内存。

GPU 环境：非强制要求；若采用深度学习方案，建议单卡显存 16GB 及以上。

存储空间：建议可用磁盘空间 50GB 及以上。

## 九、成绩评价

### （一）指标权重

本赛题的成绩评价贯穿初赛、复赛、半决赛和总决赛四个阶段。为了全面衡量算法在复杂工况下的稳定表现，将采用机器评分为主的评分方式。

### 1. 初赛阶段

数据校验与预处理权重为 50%，短周期预测指标权重为 50%。主要用于参赛者熟悉能源预测场景，掌握基本数据处理和回归预测方法。依据初赛数据集综合机器评分进行排序。初赛阶段的测评得分仅适用于初赛，用于开发验证和晋级选拔。

## 2. 复赛与半决赛阶段

短周期预测指标权重为 50%，长周期预测指标权重为 50%。依据复赛和半决赛数据集总和机器评分排序，作为总决赛的客观评分基础。

## 3. 决赛阶段

总决赛采用“70% 客观评分 + 30% 主观评分”的评价方式。其中客观评分为复赛机器测评得分标准化结果，主观评分由专家组结合技术创新性、工业可落地性、结果复现性、技术文档质量和答辩表现综合给出。

(1) 客观评分（机器测评）（权重：70%）：以测试集综合机器评分标准化结果计入。

(2) 技术创新性（权重：10%）：主要考察特征设计、模型方法和问题解决思路。

(3) 工程复现性（权重：10%）：主要考察代码运行、结果复现、效率和提交规范性。

(4) 技术文档与答辩质量（权重：10%）：主要考察文档完整性、表达逻辑、分析深度和答辩表现。

## 4. 有效成绩阈值

机器评分不低于 60 分，且不低于官方基线参考分数。

### (二) 评分细则

1. 数据一致性校验：包含多源数据时间对齐、单位统一、采样频率统一。
2. 数据预处理：包含数据完整性达标、异常值检测与修正、重复数据剔除、时间连续性处理、去噪、清洗流程与可复现。
3. 数据预测：短周期指 2 小时，长周期指 24 小时。
4. 发电优化：与预测周期保持一致；包含用于发电的煤气量。

## 十、解题思路

### (一) 知识点

本赛题主要考核工业多表数据理解、时间序列对齐、特征工程、回归建模、多模板优化、异常处理、时间切分验证、结果解释与工程复现等知识点。

### (二) 思路引导

1. 明确数据粒度，统一数据采样频率，建议统一至 15 分钟/次，并完成多个数据的时序对齐；
2. 可通过特征工程构造符合业务规则的新特征，增强模型的学习能力；

3. 发电机组基于煤气可用量进行发电，在预测未来发电量时，可先对煤气可用量进行预测。煤气可用量等于煤气发生量与消耗量之差，其中消耗量包括生产用户消耗量；

4. 建议采用时间顺序切分进行验证，避免随机切分造成未来信息泄露；

5. 建议采用滚动预测，保持输入和输出窗口长度；

6. 对于大偏差样本进行分析，识别异常工况；

7. 煤气发电计划优化的核心是优先保正常生产，其次才是降本。正常生产包括：煤气用户能够保证正常生产、气柜压力稳定（通过柜容进行判断）。此外，应尽量避免煤气放散。

### （三）注意事项

1. 严禁在特征构建或验证中使用参考时刻之后的数据；

2. 若采用深度学习方案，应注意不同字段量纲带来的归一化问题；

3. 结果文件中，预测结果不得缺失或错位。如果有发电优化方案，计划内容不得超越生产约束。

## 十一、赛题约束条件

### （一）算法约束

1. 禁止使用任何商业闭源在线推理接口或需要联网调用的模型服务；

2. 禁止通过人工逐条修正测试结果；

3. 禁止使用未来信息、测试标签泄露信息或通过测试集反推标签；

4. 进入总决赛的方案须支持离线环境完整运行。

### （二）数据使用约束

1. 参赛者仅可使用赛事官方提供的数据，严禁使用任何外部公开数据、私有数据或自行采集数据；

2. 严禁使用依赖外部数据训练得到的预训练模型权重；允许使用开源算法框架、开源代码库和通用基础依赖库，但不得借助外部数据增强模型能力；

3. 禁止传播、转售、公开展示、用于商业用途或泄露赛事数据；

4. 所有数据仅授权用于本赛事相关算法研究与评测。

## 十二、参考资源

### （一）文献资料

1. 何相君, 雷婧, 彭宝祥, 等. 钢铁副产煤气资源化利用技术及效益分析[J/OL]. 现代化工, 1-7[2026-04-25]. <https://link.cnki.net/urlid/11.2172.TQ.20260420.1511.034>.

2. 刘书含, 孙文强. “数智赋能”背景下钢铁企业副产煤气产耗量预测[C]//中国金属学会能源与热工分会, 东北大学. 第十三届全国能源与热工学术年会论文(摘要)集. 东北大学冶金学院; 2025:135136. DOI:10.26914/c.cnkihy.2025.052288.

3. XGBoost: A Scalable Tree Boosting System. KDD 2016.
4. LightGBM: A Highly Efficient Gradient Boosting Decision Tree. NeurIPS 2017.
5. CatBoost: unbiased boosting with categorical features. NeurIPS 2018.

## (二) 在线资源

### 【第九讲--常见的四种煤气发电工艺】

[https://www.bilibili.com/video/BV1fHUpBjEx6/?share\\_source=copy\\_web&vd\\_source=752a33c4774ca9d112c09c887e443e08](https://www.bilibili.com/video/BV1fHUpBjEx6/?share_source=copy_web&vd_source=752a33c4774ca9d112c09c887e443e08)

## 十三、提交要求

### (一) 初赛提交内容

参赛者需完成滚动预测，将预测结果保存为 csv 文件，其中预测模型的输入为 input.csv，短周期预测为 s\_result.csv，编码为 UTF-8，并压缩为 zip 文件后提交。

结果文件字段要求如下：

1. datetime（字符串）：时间戳格式统一为 YYYY-MM-DD HH:MM:SS。
2. input.csv（浮点型）：参赛者构建的预测模型的输入变量，包含 datetime、原始特征和特征工程构造的特征。原始特征字段名称须与赛事要求完全一致，特征工程构造的特征（如有）的字段名称需包含前缀“feat\_”。
3. s\_result.csv：必须包含 datetime，以及 generator\_1 和 generator\_all 在不同预测步长下的预测值。datetime 指滚动预测的起点时刻，其余列名（参数名）以目标时刻偏移量命名。例如，步长为 15 分钟时，列名应类似 generator\_1\_t+15\_pred, ……，generator\_1\_t+120\_pred，表示在当前起点预测未来 15 分钟到 120 分钟的 generator\_1 负荷；generator\_all 同理。预测结果不得缺行、重复，建议保留三位及以上小数，单位与原始数据保持一致。

因评测数据存在时间边界，后期滚动起点的预测目标将超出所提供的数据范围。参赛者只需基于已有历史数据正常输出预测值即可，按起点时间完整输出预测宽表，评分脚本会自动根据起点时刻与偏移量计算目标时刻，并提取真实值完成评分计算。此要求说明与复赛、半决赛、决赛相同，之后不再重复描述。

注：提交限制：每天最多提交 5 次，排行榜实时更新，以最后一次提交为准。

### (二) 复赛和半决赛提交内容

参赛者需完成滚动预测，将预测结果保存为 csv 文件，其中短周期预测结果为 s\_result.csv，长周期预测结果为 l\_result.csv，编码为 UTF-8，并压缩为 zip 文件后提交。

结果文件字段要求如下：

1. datetime（字符串）：时间戳格式统一为 YYYY-MM-DD HH:MM:SS。
2. s\_result.csv：必须包含 datetime，以及 generator\_1 和 generator\_all 在不同预测

步长下的预测值。`datetime` 指滚动预测的起点时刻，其余列名（参数名）以目标时刻偏移量命名。例如，对于步长为 15 分钟，列名应类似 `generator_1_t+15_pred`, ....., `generator_1_t+120_pred`, 表示在当前起点预测未来 15 分钟到 120 分钟的 `generator_1` 负荷；`generator_all` 同理。预测结果不得缺行、重复，建议保留三位及以上小数，单位与原始数据保持一致。

3. `l_result.csv`: 必须包含 `datetime`，以及 `generator_1` 和 `generator_all` 在不同预测步长下的预测值。`datetime` 指滚动预测的起点时刻，其余列名（参数名）以目标时刻偏移量命名，要求同 `s_result.csv`，即当设定步长为 15 分钟时，列名应类似 `generator_1_t+15_pred`, ....., `generator_1_t+1440_pred`, 表示当前器未来 15 分钟到 24 小时的 `generator_1` 负荷，`generator_all` 同理。预测结果不得缺行、重复，建议保留三位及以上小数，单位与原始数据保持一致。

### （三）决赛提交内容

1. 算法代码：包含数据读取、预处理、特征构建、模型训练、预测推理与优化等完整代码。

2. 模型文件：训练好的模型参数文件。

3. 技术报告：PDF 或 Word 格式。

4. README 文件：说明运行环节、依赖安装、运行命令和输入输出目录结构。

5. `s_result.csv`: 短周期预测结果，单位与原始数据保持一致。

6. `l_result.csv`: 长周期预测结果，单位与原始数据保持一致。

7. `opt_result.csv`: 优化结果。优化结果包含发电用的高炉、焦炉、转炉煤气字段的浮点型数据，字段需包含前缀“`opt_`”，单位与原始数据保持一致。

CSV 文件应使用 UTF-8 编码，字段名称须与赛事要求完全一致；每个样本编号只能出现一次；不得缺行、重复或附加多余字段。预测对象必须包含煤气发电量（发电机组负荷），煤气发生量和消耗量作为可选对象。优化对象必须包含用于发电的高炉、焦炉、转炉煤气量，其余作为可选对象。预测值建议保留三位及以上小数。

如赛事组织方需要，可额外提交视频、答辩 PPT 和补充说明材料。

### （四）提交规范

1. 所有提交材料须保证可读取、可解压、可运行；

2. 初赛压缩包命名规则为 `teamname_gas_predict_prelim.zip`，复赛压缩包命名规则为 `teamname_gas_predict_final.zip`；

3. 若代码与结果无法复现，赛事方有权取消成绩或奖项；

4. 最终解释权归赛事主办方所有。

## 十四、奖励设置

为了鼓励参赛选手参赛积极性，激发在复杂制造场景下的技术创新潜能，本赛题

不仅设立了丰厚的现金奖励，更针对优秀的青年算法人才开辟了专属的职业发展绿色通道。具体激励设置如下：

### （一）赛事奖金设置

本赛题将根据总决赛的最终综合成绩（含线上客观评测与线下答辩），对全国总决赛一等奖前六名参赛团队按照如下标准颁发奖金：

1. 冠军奖：第 1 名，奖金 8000 元/每团队
2. 亚军奖：第 2-3 名，奖金 6500 元/每团队
3. 季军奖：第 4-6 名，奖金 3000 元/每团队

### （二） 人才招募专属特权

为加速“人工智能+”创新成果向新质生产力的转化，切实打通产学研用链路，针对在本赛题中展现出卓越代码能力与工程落地思维的优秀参赛选手（不限于前六名获奖团队成员），出题方还将额外提供以下极具含金量的职业发展激励：

#### 1. 核心岗位面试直通车

在复赛或总决赛中表现突出的核心算法选手，将获得人工智能研究院算法研发岗位的“校招面试直通车”。凭此卡可直接免除在线笔试与简历初筛环节，直通终审技术面，优先锁定高潜算法人才核心席位。

#### 2. 专项精英实习计划

定向开放工业计算机视觉、大模型前沿应用等核心研发方向的优质实习岗位。受邀实习生将直接参与真实的高级别产线视觉项目，享有充足的算力资源支持与资深算法研究员的 1v1 业务指导，积累宝贵的工业界实战经验，实习期间表现优异者可获全职 Offer 提前转正机会。

## 十五、其他说明

**公平性：** 严禁任何形式的作弊行为，包括但不限于数据泄露、模型预训练数据与测试数据重叠、抄袭他人代码等。一经发现，立即取消参赛资格，并追究相关责任。

**知识产权：** 参赛者提交的作品必须为原创，未在其他比赛中获奖或公开发表。比赛主办方有权对参赛作品进行展示、宣传等相关活动，但知识产权仍归参赛者所有。

## 十六、联系方式

赛题交流 QQ 群：1093810052

邮 箱：huanglinya@njsteel.com.cn

报名官网：www.aicomp.cn