**2020 CCF BDCI 路况预测赛题**

**赛题背景**  
移动互联网时代的到来让所有移动设备的持有者都可以成为道路通行能力的描绘者, 滴滴平台收集了海量的高质量司乘轨迹数据, 可以对实时道路拥堵状况有良好的建模能力. 如果可以基于实时和历史的路况信息, 对未来的路况状态有较精准的预估, 无疑对出行决策, 缓解城市拥堵等场景有至关重要的作用. 然而, 未来的路况预估仍然是十分困难的, 未来路况会受到时间周期, 道路通行能力, 路网上下游拓扑, 导航流量以及道路突然状况等多种因素的影响. 此次竞赛诚邀参赛者基于滴滴提供的实时与历史路况状态信息以及道路属性等信息, 精准预估未来某时间段内的路况状态, 助力城市规划与智能出行方案.  
  
**赛题任务**  
根据滴滴提供的道路小段的实时和历史路况状态特征, 道路基本属性以及路网拓扑关系图, 预测未来一段时间内道路小段的路况状态(即畅通, 缓行和拥堵几类状态).  
  
**数据说明**  
本次比赛提供滴滴平台2019年7月1日至2019年7月30日西安的实时和历史路况信息, 以及西安市的道路属性和路网拓扑信息.  
  
**术语解释**  
**link:**对完整道路按照拓扑切分后得到的小段, 由唯一id标识. 出于数据安全考虑, 隐去了经纬度等真实地理信息.  
**路况状态:** 根据道路的平均车速, 道路等级等信息对道路通行状态的描述, 分为畅通, 缓行, 拥堵三种状态, 分别对应滴滴地图展示的绿色, 黄色, 红色.  
**时间片:** 对时间的离散化描述. 一般以2分钟为一个单位. 2分钟内认为道路的路况状态是统一的.

##### **数据一:历史与实时路况**

  总体格式：linkidlabelcurrent\_slice\_idfuture\_slice\_id;recent\_feature;history\_feature

|  |  |
| --- | --- |
| 字段名称 | 字段含义 |
| link | 小段的id |
| label | 对应时间的link的路况状态 |
| current\_slice\_id | 当前时间片id |
| future\_slice\_id | 待预测时间片id |
| recent\_feature | 近期n个时间片路况特征，n=5，时间片之间空格分隔，字段之间,分隔. 具体格式：时间片:路况速度,eta速度,路况状态,参与路况计算的车辆数. 特征都为0时，说明此时间片无车经过 |
| history\_feature | 历史同期n个时间片路况特征，星期之间;分隔，共4组（-28,-21,-14,-7），每组格式和recent\_feature一致 |

##### 数据二:道路属性

数据示例：

|  |  |  |
| --- | --- | --- |
| 字段名称 | 字段类型 | 字段含义 |
| linkid | categoric | link的id |
| length | numeric | link的长度，以m为单位 |
| direction | categoric | link的通行方向 |
| pathclass | categoric | link的的功能等级 |
| speedclass | categoric | link的速度限制等级 |
| LaneNum | categoric | link的车道数 |
| speedlimit | numeric | link的限速，以m/s为单位 |
| level | categoric | link的level |
| width | numeric | link的宽度，以m为单位 |

##### 数据三:路网拓扑

|  |  |
| --- | --- |
| key | value |
| link | 下游linkid1,下游linkid2,下游linkid3,... |

**城市交通指数**

2018年深圳、成都、西安、苏州、济南、海口市各城市交通指数数据，包含城市级、辖区级、道路级的交通指数数据（Travel Time Index ,TTI）和平均行驶速度。

开放城市：深圳、成都、西安、苏州、济南、海口

开放时间： 2018年1月1日-2018年12月31日

数据内容：各城市级、辖区级、道路级滴滴基于浮动车轨迹的城市交通指数计算服务所得交通指数数据和平均行驶速度。

数据量：14G(压缩前)

数据格式：压缩包内共三个文件，city\_district.txt、boundary.txt、road.txt。

city\_district.txt为城市和辖区的TTI数据，road.txt为道路的TTI数据，boundary.txt为TTI对象的几何范围，用WKT1描述。

使用前请阅读readME.txt文件

立即申请

滴滴基于浮动车轨迹的城市交通指数计算服务

本数据基于滴滴开源算法-基于浮动车轨迹的城市交通指数计算服务，开源代码请访问：<https://github.com/didi/TrafficIndex>

著作权及引用格式

如果申请人对数据进行的任何分析挖掘而产出的论文、行业报告等，需注明数据出处：[https://gaia.didichuxing.com](https://gaia.didichuxing.com/) ，并在致谢处署名：数据来自滴滴出行“盖亚”数据开放计划，英文（Data source: Didi Chuxing GAIA Initiative）。

city\_district,road文件说明

| 字段 | 说明 | 样例数据 |
| --- | --- | --- |

|  |  |  |
| --- | --- | --- |
| obj\_id | TTI对象ID | 841 |
| batch\_time | 时间 | 2018-01-01 00:00:00 |
| tti | 交通指数数据 | 1.18665 |
| speed | 平均速度 | 47.3983 |

boundary文件说明

| 字段 | 说明 | 样例数据 |
| --- | --- | --- |

|  |  |  |
| --- | --- | --- |
| obj\_id | TTI对象ID | 283203 |
| obj\_name | TTI对象名称 | 百日红中路:三环路,银杏大道 |
| geom | TTI对象几何范围 | MULTILINESTRING((104.13743 30.60591,104.13832 30.60538),(104.13832 30.60538,104.13843 30.60532),(104.13843 30.60532,104.13855 30.60525),(104.13855 30.60525,104.13921 30.6049,104.13984 30.60454)…) |

数据采集和处理方法

1. 本数据开放的目的仅为了用于学术研究。  
2. 其中boundary文件中对道路描述经过抽稀化处理以供研究使用，用WKT1描述。  
注：1）WKT 说明： <https://www.opengeospatial.org/standards/wkt-crs>  
3. Speed速度计算基本思想：如果一条 link 有两个连续时间片，分别为 t1、t2，link 长度为S，那么 t1 到 t2 这段时间内，link 的平均速度 v 为 v = 2·S / (t1 + t2)。  
其中，link为地图上一条固定长度、有方向的路段，矢量数据。TTI指数计算基本思想：在同一条link在一个时间片内，TTI = 自由流速度 / 实际速度。

集合s = { Link1 ,Link2 ,Link3 ,Link4 , … …LinkN }

TTI =

∑Ni=1  LiVi·Wi

∑Ni=1   LiVfree\_i·Wi

speed =

∑Ni=1   Li·Wi

∑Ni=1  LiVi·Wi

集合中link的总数为N，Li 为link的长度，Wi 为link的权重，Vfree\_i 为link的自由流速度，Vi 为link的实时路况速度。  
关于自由流速度和权重的计算方式，请参考： <https://github.com/didi/TrafficIndex>

已申请